Genetic divergence, reproductive isolation and the early stages of speciation

Silje Hogner

Dissertation presented for the degree of Philosophiae Doctor

Natural History Museum

Faculty of Mathematics and Natural Sciences

University of Oslo

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1305

ISSN 1501-7710

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.
The thesis is produced by Akademika publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Supervisor

Arild Johnsen
Associate Professor
Natural History Museum
University of Oslo
Norway

Co-supervisor

Jan T. Lifjeld
Professor
Natural History Museum
University of Oslo
Norway

Adjudicating committee

1st opponent
Jacob Höglund
Professor
Department of Ecology and Genetics
Population Biology and Conservation Biology
Uppsala University
Sweden

2nd opponent
Laura Kvist
Docent/Phd
Department of Biology
University of Oulu
Finland

Administrative leader
Øystein Wiig
Professor
Natural History Museum
University of Oslo
Norway

Acknowledgements

The work of this PhD thesis was undertaken at the Natural History Museum, University of Oslo, and would not have been possible to complete without the help of several people. First, I would like to thank my supervisor Arild Johnsen. It has been a privilege to have you as my supervisor and to work with you. Your enthusiasm, knowledge and support have been a great inspiration. Many thanks also to Jan T. Lifjeld for being my co-supervisor. To both of you, thank you for inspiring me to be a better scientist and for always taking the time to discuss my work.

I would like to thank Terje Laskemoen, Gunnhild M. Marthinsen, Lars Erik Johannessen, Melissah Rowe, Jostein Gohli, Even Stensrud, Oddmund Kleven, Trond Øigarden, Becky Cramer, Taiwo Crossby Omotoriogun and Jarl Andreas Anmarkrud, also known as the "bird group". Thank you for providing such a friendly, dynamic and supportive research environment. I have enjoyed the numerous discussions we have had during coffee-breaks and lunches. A special thanks to Terje, for help with sperm analyses and statistics. Lars Erik thank you for helping me with all the formatting, and Gunnhild thank you for valuable help in the lab.

Melissah, thank you for helping me improve my writing, for all the fruitful and helpful discussions we had during the final stages of my PhD and for always encouraging me to continue. It was a pleasure sharing an office with you.

Kjersti S. Kvie, thank you for being a great friend and a great colleague. Thank you for always having time for conversation and support when things may have seemed too chaotic and unmanageable.

Eirik Rindal, thank you for valuable help with statistics. Thanks also to Vegar Bakkestuen for help with mapping of geographical distributions. I am also grateful to all my colleagues at NHM for making my PhD a memorable time.

This thesis would not have been possible to carry out without field work and extensive sampling of both blood- and sperm samples. Thanks to all who were involved in the sampling of both birds and moths and a special thanks to all of you making it possible for me to do field work in Trysil, Karmøy, Heimdalen, Bulgaria, France and Czech Republic. I had a great time collecting samples in all of those places, and it will not be forgotten.

I also want thank my supervisor during my master thesis, Glenn-Peter Sætre. Thank you for helping me with the flycatcher paper, for making it possible for me to have a second field trip to Bulgaria catching semicollared flycatcher and for inspiring me to start working on a PhD.

Tore Oldeide Elgvin and Stein Are Sæther, thank you for all the discussions we had when I was struggling with the flycatcher paper. Jo Skeie Hermansen, thank you for taking time off your PhD to help me catch birds in Bulgaria.

To all my co-authors (Jiri Porkert, Tamer Albayrak, Bekir Kabasakal, Leif Aarvik, Václav Pavel, Bohumír Chutný, Javier García, Marie-Christine Eybert, Ekaterina Matsyna, Thomas Borge and Torbjørn Bruvik), thank you for sharing your data with me and for helping me improve our manuscripts.

Marius and Natalia, thank you for help taking care of my horse, especially during the last final weeks of my writing process. Keeping him would not have been possible without your help.

Finally a big thanks to my family. Eirik, thank you for letting me use your apartment while you were working abroad. Mom and dad, thank you for always believing in me, never giving up and always encouraging me to continue. Thank you for taking interest in my work, no matter how abstract it might seem to you.

Oslo, December 2012
Silje Hogner

Contents

List of papers 1
1 Abstract 2
2 General introduction 4
2.1 Main aims 7
3 General methods 8
3.1 Study taxa 8
3.2 Field work and sampling 12
3.3 DNA extraction, PCR and sequencing 13
3.4 Sperm analysis 13
4 Summary of publications 14
4.1 Paper I: Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus* 14
4.2 Paper II: Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*) 15
4.3 Paper III: Rapid sperm evolution in the bluethroat (*Luscinia svecica*) subspecies complex 17
4.4 Paper IV: Increased divergence but reduced variation on the Z chromosome relative to autosomes in *Ficedula* flycatchers: differential introgression or the faster-Z effect? 17
5 General discussion 19
5.1 Speciation in reverse 19
5.2 Early speciation/rapid evolution of phenotypic traits 20
5.3 Recent speciation, Z chromosome evolution and the faster-Z hypothesis 21
5.4 DNA barcoding 22
5.5 Genetic divergence and the process of speciation 23
5.6 Speciation and sexual selection 24
6 Conclusions 26
7 Future prospects 27
8 Literature cited 29

Individual papers I-IV

List of papers

I Silje Hogner, Terje Laskemoen, Jan T. Lifjeld, Jiri Porkert, Oddmund Kleven, Tamer Albayrak, Bekir Kabasakal and Arild Johnsen (2012). Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. *Ecology and Evolution* 2: 2974–2988.

II Kjersti S. Kvie, Silje Hogner, Leif Aarvik, Jan T. Lifjeld, and Arild Johnsen (2013). Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*). *Ecology and Evolution* 3: 126–144.

III Silje Hogner, Terje Laskemoen, Jan T. Lifjeld, Václav Pavel, Bohumír Chutný, Javier García, Marie-Christine Eybert, Ekaterina Matsyna and Arild Johnsen. Rapid sperm evolution in the bluethroat (*Luscinia svecica*) subspecies complex. Manuscript to be submitted to Behavioral Ecology and Sociobiology.

IV Silje Hogner, Stein A. Sæther, Thomas Borge, Torbjørn Bruvik, Arild Johnsen and Glenn-Peter Sætre (2012). Increased divergence but reduced variation on the Z chromosome relative to autosomes in *Ficedula* flycatchers: differential introgression or the faster-Z effect? *Ecology and Evolution* 2: 379–396.

1 Abstract

The process of speciation is the splitting of single populations into two or more distinct, reproductively isolated taxa. Common modes of speciation are sympatric, allopatric and parapatric speciation, with speciation in allopatry being the most frequently documented mode to date. In allopatric speciation, geographical barriers physically separate populations, allowing these now isolated groups to evolve reproductive barriers, i.e. barriers to successful reproduction, which can take the form of premating, postmating prezygotic or postzygotic barriers. Species level phylogenies derived from molecular data may provide an indirect record of speciation events, and can, when combined with morphological traits, be used to investigate at what stage in the speciation process (e.g. early speciation, recent speciation, reversed speciation) taxa currently are. In this thesis, I used a range of molecular methods and morphological analysis to investigate different stages in the speciation process. More specifically, I investigated four different species/species complexes exhibiting varying degrees of genetic and morphological divergence in order to investigate where in the speciation process taxa are and to discuss the evolutionary processes involved in the speciation events.

First, the phylogeographic pattern of the common redstart (*Phoenicurus phoenicurus*) was described and the level of genetic divergence quantified. In this system, high divergence within the mitochondrial DNA (5% K2P distance, COI) combined with low morphological divergence appears to reflect reversed speciation. Second, I found a similar pattern of high genetic divergence (1.5-4.1% K2P distance, COI) in the autumnal moth (*Epirrita autumnata*), for which low morphological divergences have previously been found. Moreover, an association between the moths' mtDNA divergence and infection by different Wolbachia strains was found, and I suggest that this association maintains the mitochondrial variation. In contrast to these two studies, the bluethroat (*Luscinia svecica*) subspecies complex was characterized by exhibiting low genetic divergence (mean genetic distance 0.7%, K2P distance, COI) and high morphological differences and, as such, appears to exhibit signs of early speciation. Importantly, these contrasting patterns may be explained by differences in both ecology and sexual selection pressures experienced by each of the species/populations, with the bluethroats being subject to strong diversifying sexual selection for male primary and secondary sexual characters.

A third goal of this thesis was to investigate whether sperm characters and genetic markers evolve at different speeds. In the bluethroat subspecies complex, where mitochondrial divergence was low, I found evidence of rapid evolution of sperm morphology, suggesting that rapid evolution of gametes may be an important factor involved in the early stages of speciation. Finally, I studied the black-and-white *Ficedula* flycatchers, a group of species suggested to have undergone recent speciation, in order to investigate variation in the rate of evolution between the Z chromosome (i.e. sex chromosome) and the autosomes. In this system, I found contrasting patterns in the evolution of the Z chromosome versus the autosomes. Specifically, my results revealed increased divergence and reduced variation on the Z chromosome compared to the autosomes, a finding that is best explained by the faster-Z hypothesis. As the Z chromosome has been linked to sexually selected traits in the *Ficedula* flycatchers, I suggest the contrasting pattern of evolution on the Z vs. autosome may have implications for the process of speciation processes in these species. In conclusion, my thesis highlights the utility of combining patterns of genetic and phenotypic divergence to identify at what stage of the speciation process taxa occur and how variation in evolutionary rates between traits can contribute to our understanding of the speciation process.

2 General introduction

Speciation is the process in which a single population splits into two or more distinct, reproductively isolated taxa (Mayr, 1963). Understanding the processes underlying speciation remains a fundamental challenge in biology. Moreover, understanding the mechanisms that generate species diversity is essential if we are to conserve biological diversity at both the local and global scale. In this thesis, I make use of four study organisms/groups that show varying degrees of morphological and genetic divergence to investigate early stages of the speciation process and factors that may characterize each of these stages and contribute to speciation.

Common modes of speciation are allopatric, parapatric and sympatric speciation (Coyne & Orr, 2004). Allopatric speciation occurs when populations are geographically isolated by an extrinsic barrier (e.g. mountain range or body of water) and reproductive isolating mechanisms (i.e. reproductive barriers) have sufficient time to evolve such that, if populations come into secondary contact, individuals are no longer able to interbreed. Under such a scenario, the two populations are considered distinct species (Mayr, 1963). In contrast, parapatric speciation occurs when two divergent populations are only partially separated, and while individuals from these populations may occasionally come into contact and reproduce, selection for specific behaviors or isolating mechanisms eventually prevents them from interbreeding (Endler, 1977). Finally, sympatric speciation occurs when two or more species evolve from a single ancestral species while inhabiting the same geographic area (Maynard Smith, 1966; Bush, 1994). Evidence suggests that the majority of speciation occurs in allopatry (Coyne & Orr, 2004), indicating that geography is an important driver of species divergence and that gene flow between populations tends to restrict the speciation process.

Sexual selection arises from differences in reproductive success among individuals within a population (Darwin, 1871). Such differences may occur if there is choice for specific traits in one sex by the other (e.g. female choice) or through competition between members of the same sex (e.g. male-male contest competition, sperm competition; Andersson, 1994). Sexual selection can be a powerful evolutionary force and is suggested to increase the rate of reproductive divergence between populations and thereby drive the evolutionary diversification of clades (Schluter & Price, 1993; Barraclough *et al.*, 1995; Panhuis *et al.*, 2001). For example, males may evolve secondary sexual signals (i.e. traits

that reveal information about the bearer to females) through female mate preferences, which may then create barriers to gene flow and hence promote reproductive isolation (Barraclough *et al.*, 1995), which is critical to the process of speciation. Reproductive barriers fall into three major categories: premating, postmating prezygotic and postzygotic. Premating barriers prevent or reduce the likelihood of copulation and the subsequent formation of hybrid zygotes, and may arise from behavioral or ecological (i.e. differences in species' behavior or ecology that prevent courtship or copulation) or mechanical (i.e. inhibition of normal copulation due to incompatibility of reproductive structures) isolation. Postmating prezygotic barriers act after the transfer of gametes (i.e. sperm or pollen) but prior to fertilization, and may arise as copulatory behavior isolation (i.e. behavior of one individual is insufficient to allow normal fertilization) or as gametic isolation (i.e. gamete is transferred but unable to fertilize an egg). Finally, postzygotic barriers act after fertilization, manifesting as hybrid inviability (i.e. hybrids suffer developmental problems causing full or partial lethality) or sterility (i.e. hybrids are partially or completely sterile; Coyne & Orr, 2004).

Species level phylogenies derived from molecular data provide an indirect record of events leading to the diversification of taxa. Phylogenies describe the evolutionary history of a species or higher taxonomic unit, especially in reference to lines of descent and relationships among groups of organisms. A related field of biology is phylogeography, which concerns the geographical mapping of neutral genetic structure within and among closely related species (Avise, 2000). Importantly, phylogeography can sometimes reveal major historical lineages and historical changes in population size and range (Avise *et al.*, 1987; Avise, 2000). Thus phylogenies and the field of phylogeography can be powerful tools that help us understand the process of speciation. For example, deep DNA divergence in species suggests that periods of allopatry have played an important role in the process of speciation. More specifically, the amount of genetic divergence between populations acts as a measure of time since separation (Hewitt, 2004).

Traditionally, phylogeographic studies have utilized mitochondrial markers for investigating relationships at or below the species level. They have done so for a number of reasons. First, mitochondrial DNA has a relatively high mutation rate relative to nuclear DNA; though actual rates may vary among markers. Additionally, due to predominantly maternal inheritance, they exhibit little or no recombination (Avise, 2000; but see Kvist *et al.*, 2003). Relative to nuclear introns, mtDNA has a smaller effective population size, faster coalescent time, more rapid evolution at the nucleotide sequence level and extensive

intraspecific polymorphisms (Avise, 2000). Thus mitochondrial markers allow distinct taxonomic units to be easily identified and discriminated from one another (Avise, 2000; Hewitt, 2001; Zink & Barrowclough, 2008). Furthermore, mitochondria are present in the majority of cells, occur in high copy numbers and are relatively easy, rapid and cheap to sequence (Zink & Barrowclough, 2008). Consequently, when a species is well sampled throughout its geographical range, a phylogenetic tree of mtDNA haplotypes, rooted with an outgroup, can reveal whether closely related haplotypes coexist throughout the entire range or whether some/all haplotypes show localized distributions within the larger range (Zink & Barrowclough, 2008). This has been the essence of phylogeography (Avise *et al.*, 1987) and knowledge of the geographic distribution of haplotypes continues to be a vital component of modern phylogeography. Nevertheless, phylogeography has been criticized for relying too heavily upon this single gene system (i.e. mitochondrial DNA) as a means to determine evolutionary decent (e.g. Ballard & Whitlock, 2004; Edwards *et al.*, 2005; Bazin *et al.*, 2006). Moreover, the use of a single gene system has been criticized because selection pressures may vary across regions and thus the evolutionary patterns observed in mitochondrial DNA may not be representative of the evolutionary history of the entire genome. Furthermore, the possibility of amplifying pseudogenes and interspecific hybridization can obscure the delineation of lineages. However, these pitfalls can often be avoided through the use of molecular and numerical analyses and by testing for congruence between nuclear and mitochondrial genes (Bermingham & Moritz, 1998).

An alternative to mitochondrial markers is the use of nuclear introns, which are non-coding DNA regions situated between coding exon regions (Gilbert, 1978). Intron sequences can, in theory, be treated in an identical manner to mtDNA sequences to construct gene trees utilizing individual haplotypes as terminal taxa in phylogenetic analyses (Zink & Barrowclough, 2008). However, because introns typically have slower rates of mutation relative to mtDNA (Willows-Munro *et al.*, 2005), they tend to display less variation per sequenced base. This, together with the four-fold higher effective population size compared to mitochondrial genes, means that mtDNA is able to detect more recent splits that nuclear loci are unable to resolve (Zink & Barrowclough, 2008). Among nuclear markers, Z-linked and autosomal loci have different divergence times due to differences in their population sizes; in species with female heterogamy (e.g. birds, moths), the effective population size of Z-linked loci is (ideally) 3:4 of the autosomal loci (due to females only having one copy of the Z chromosome). Therefore, under conditions of a balanced sex ratio and equal mutation rates, the neutral expectation is that the

nucleotide variation of Z-linked genes would be 3:4 that of autosomal variation (Ellegren, 2009). However, several bird studies have reported this ratio to be below 3:4, suggesting that additional forces may reduce variation on the Z chromosome relative to the autosomes (Berlin & Ellegren, 2004; Borge *et al.*, 2005b; Backström *et al.*, 2010; Storchova *et al.*, 2010).

2.1 Main aims

In this thesis I investigated four study systems that show varying degree of divergence in mitochondrial DNA. First, I used two species (common redstart, autumnal moth) for which high mitochondrial divergence has been found within sympatric populations (Johnsen *et al.* 2010, Johnsen, Lifjeld & Aarvik unpublished data) in order to investigate what stage of speciation these species currently occupy (paper I and II). Next, I used a group of allopatric bluethroat subspecies exhibiting low mitochondrial divergence, to test if there is a relationship between genetic divergence and sperm divergence across subspecies (paper III). Finally, I studied a group of four allopatric sister species (flycatchers) that exhibit intermediate levels of mitochondrial divergence (relative to the autumnal moth and the common redstart; Sætre *et al.* 2001). Additionally, in this group, a contrasting pattern between Z-linked and autosomal loci has been observed in two of the four species (Borge *et al.*, 2005b), and here I investigated whether or not this pattern holds for the complete species complex (paper IV).

I chose to work with these study groups because of the contrasting patterns of divergence and geographical distribution. More generally, these groups allowed me to investigate if differences in mitochondrial divergence between species could be explained by these species being in different stages of speciation (e.g. speciation in reverse [paper I and II], early speciation [paper III], recent speciation [paper IV]). Finally, these study systems also allowed me to investigate whether morphological and molecular traits vary with respect to speed of evolution, and to consider how variation in evolutionary rates might be associated with the speciation process (paper III and IV).

3 General methods

3.1 Study taxa

The common redstart (*Phoenicurus phoenicurus*)

The common redstart is a small (~15g) sexually dimorphic passerine bird belonging to the family Muscicapidae (Figure 1). Distributed across Europe, Asia and North Africa, the species breeds in the Western Paleartic and winters in North Africa (Cramp, 1988b). The breeding system of the common redstart is predominantly characterized as social monogamy (Kleven *et al.*, 2007), though instances of polygyny (males mating with several females in different territories) have been observed (del Hoyo *et al.*, 2005). Based on differences in the male plumage coloration, two subspecies of the common redstart have been described: *P. p. phoenicurus* and *P.p. samamisicus* (Cramp, 1988b). Moreover, deep mtDNA divergence have been observed using cytochrome c oxidase subunit I (COI, ~ 5%) in sympatric populations (Johnsen *et al.*, 2010), and studies have shown that this species exhibit low sperm competition (Kleven *et al.*, 2007).

Figure 1: Male common redstart captured in Trysil, Norway. Photo: Silje Hogner.

The bluethroat (*Luscinia svecica*)

The bluethroat is a small (~18g), sexually dimorphic passerine bird belonging to the family Muscicapidae (Figure 2). The breeding range of the bluethroat extends from the western Palearctic to eastern Eurasia, whereas the wintering grounds are restricted to Africa and Southeast Asia (Cramp, 1988a). The bluethroat subspecies complex is putatively comprised of 10 subspecies, all of which are migratory, though migration distances vary among subspecies. Males possess a colorful throat patch which they display during courtship (Figure 2; Peiponen, 1960). In contrast, throat coloration is absent or highly reduced in females (Johnsen *et al.*, 2006). Importantly, the color of this throat patch varies among subspecies. In the five subspecies used in the present study, the throat patch is chestnut-colored in the nominate, *L. s. svecica*, white in *L. s. cyanecula,* (Central Europe) and *L. s. namnetum* (Atlantic coast, France), mostly absent in *L. s. azuricollis* (Spain), and white or chestnut in the *L. s. volgae* (Russia) (Cramp, 1988a). In addition, male size varies across the subspecies, with the *L. s. namnetum* males being the smallest of all subspecies and *L. s. magna* the largest (Cramp, 1988a; Johnsen *et al.*, 2006). Moderate support for subspecies classification has been found using mitochondrial markers (Questiau *et al.* 1998; Zink *et al.* 2003). Specifically, Questiau *et al.* (1998) found differences between *L. s. svecica* and *L. s. namnetum*, while Zink *et al*. (2003) in a study of seven subspecies found support for two clusters, one northern clade, and one more southern clade of central and southern European subspecies. Additionally, Johnsen *et al*. (2006) found evidence of genetic differentiation between the following five subspecies, *L. s. svecica*, *L. s. cyanecula*, *L. s. namnetum, L. s. magna* and *L. s. azuricollis* based on microsatellite data. Finally, sperm competition has been shown to be high in this species (Krokene *et al.*, 1996; Johnsen & Lifjeld, 2003).

Figure 2: Male bluethroat displaying in Heimdalen, Norway. Photo: Bjørn A. Bjerke.

The European black-and white *Ficedula* flyctachers

The European black-and-white flycatcher complex consists of four species; pied (*F. hypoleuca*), collared (*F. albicollis*), semicollared (*F. semitorquata*) and Atlas (*F. speculigera*), all of which are small (~13g), sexually dimorphic passerines belonging to the family Muscicapidae (Figure 1, paper IV). These species breed in forested areas in North Africa, Europe and the near East during spring and summer, and spend the rest of the year in tropical habitats in Africa (Lundberg & Alatalo, 1992). Phylogeographic evidence suggests that these taxa are genetically distinct species that originated from a single ancestor in allopatry around the Mediterranean Sea during Pleistocene (Sætre *et al.*, 2001). Genetic evidence based on mitochondrial markers (ND6, cytochrome b, partial RNA genes) suggests that these four lineages diverged approximately 1.5-2 million years ago (Sætre *et al.*, 2001). Consistent with these estimates, COI data suggests a divergence time of around 1-1.5 million years ago (Figure 3; this thesis). Different genetic and morphological markers have been used to investigate these species, especially the pied and collared flycatcher, where a contrasting pattern of increased divergence and reduced variation are found on Z chromosomes compared to autosomes (Borge *et al.*, 2005b; Ellegren *et al.*, 2012). Finally, the pied flycatcher has been shown to have low to moderate levels of sperm competition (e.g. Lifjeld *et al.*, 1991; Rätti *et al.*, 1995), while the collared flycatcher exhibits moderate to high levels of sperm competition (Sheldon & Ellegren, 1999; Krist *et al.*, 2005).

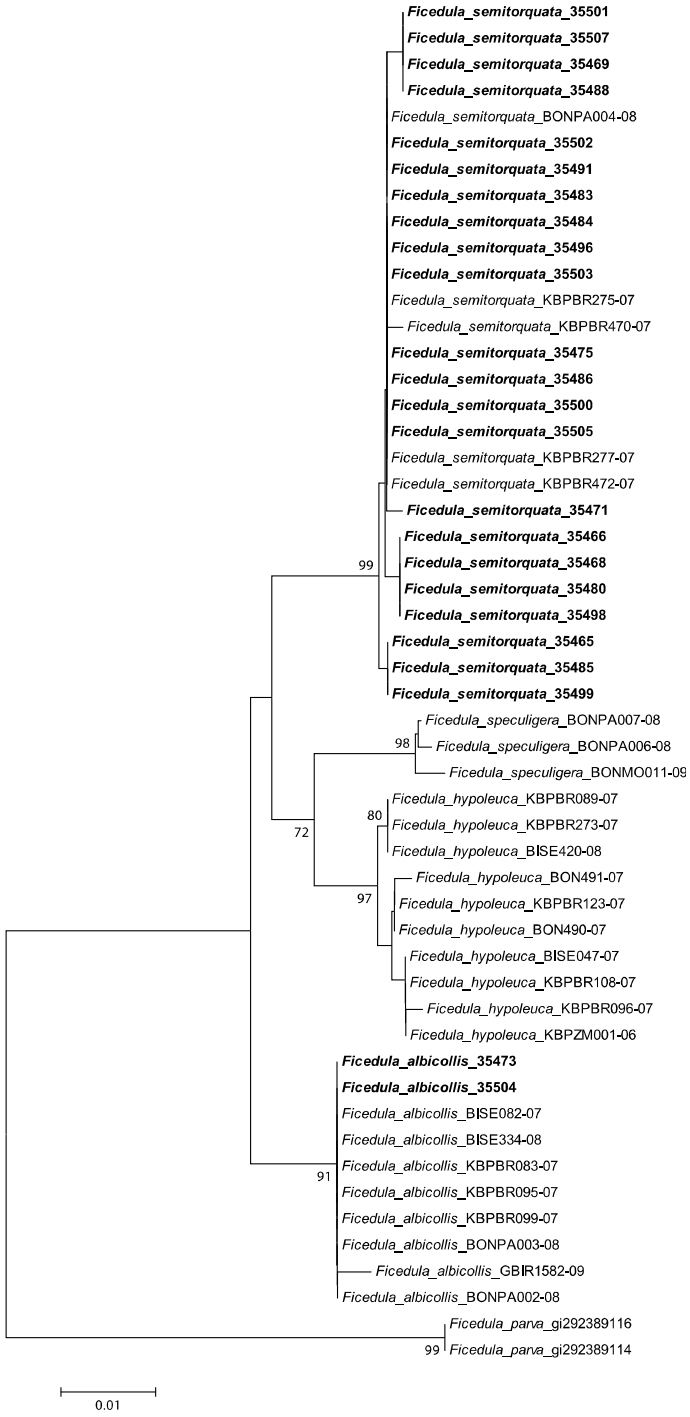


Figure 3: A Neighbour-joining analysis of 52 *Ficedula* flycatchers (26 from BOLD, 2 from Genbank, 24 unpublished) based on COI (Kimura 2 Parameter substitution model). Bootstrap support (10000 iterations) is shown at each node. Unpublished individuals marked in bold.

The autumnal moth (*Epirrita autumnata*)

The autumnal moth belongs to the family Geometridae (Lepidoptera; Figure 1, paper II). This species is found across Europe, throughout the Caucasus and east to Mongolia and Japan. Larvae feed on a wide variety of deciduous trees, including birch (*Betula*), alder (*Alnus*) and willow (*Salix*) (Aarvik *et al.*, 2009), and populations regularly undergo size fluctuations according to a 9-10 year cycle. The species is considered of agricultural importance because, during periods of high larvae densities, larval feeding results in significant damage and defoliation to mountain birch forests (Yang *et al.*, 2008). Importantly, identification of *E. autumnata* is challenging due to high within-species morphological diversity, especially in terms of pigmentation and size, and because of close resemblance to sister taxa such as *E. dilutata* and *E. christyi* (Hausmann & Viidalepp, 2012). Preliminary results from DNA barcoding of Scandinavian moths and butterflies (Lepidoptera) have revealed a discrepancy between present species delineation and levels of sequence divergence (up to 4 %, COI) in the genus *Epirrita* (Johnsen, Aarvik & Lifjeld, unpublished data), which suggest that this group might consist of several cryptic species.

3.2 Field work and sampling

Samples were obtained from museum collections and collaborators (mainly from Europe, but also from Asia and North Africa) or by sampling of new material from wild populations. Specifically, I collected fresh samples from the following populations: redstarts were sampled during 2009-2011 in Trysil, Norway (61º14'N, 12º17'E) and Hradec Kràlové, Czech Republic (50°10'N, 15°56'E). Bluethroats were sampled during 2009-2010 in Heimdalen, Norway (61º25'N, 08 º52'E) and 2011 in Briere (47º21'37.9"N, 2º12'5.3"W) and Guérande (47º20'N, 2º25'W), France. Semicollared flycatchers were sampled from Kamcheya, Bulgaria (42º53'N, 26º58'E) during 2007 and 2011, and Atlas flycatchers were sampled from Azrou, Morocco (33º26'N, 5º13'W) during 2008. All field work was conducted during the peak of the breeding season for each population. Further details of species localities are provided in the individual papers.

Birds were trapped using mist nets, bait- and box traps, with or without the use of vocal playback. Blood was sampled via puncture of the brachial vein on the right wing. For all individuals ~ 25 µL of blood was collected and stored in either Queens Lysis Buffer (Seutin *et al.*, 1991) or ethanol. For bluethroats and redstarts sperm samples were collected via cloacal massage (Wolfson, 1952; Kleven *et al.*, 2008). All birds were released

unharmed into their home territories. Moths were collected using light traps, placed in jars and returned to the laboratory for processing.

All bird species used in this thesis are good study species due to their relative robustness; handling of individuals for blood sampling and morphological measurements does not seem to cause individuals stress or affect their breeding success. Moreover, populations of both the *Ficedula* flycatchers and the common redstart breed readily in artificial nest boxes. All necessary permits to catch and sample birds were obtained prior to each field season. For the moths, no permits were required for collection of individuals. In addition, these moths often appear in high numbers making them easy to locate and sample.

3.3 DNA extraction, PCR and sequencing

DNA was extracted from tissue samples (i.e. blood, skin, muscle, middle leg of butterflies) following standard protocols using commercially available extraction kits (Omega Bio-tek, Georgia, USA; Mole Genetics AS, Lysaker, Norway; Qiagen AB, Sollentuna, Sweden). Both fresh and museum samples were analyzed using identical techniques, though a variety of molecular markers were utilized. In all instances, polymerase chain reactions (PCR) was used to amplify DNA fragments (see individual papers for further details of primer and PCR conditions) before they were sequenced using the Sanger method (Sanger *et al.*, 1977).

3.4 Sperm analysis

Following collection, fresh sperm samples were immediately diluted in a small volume of phosphate buffered saline and then fixed in 5% formalin solution. To examine sperm morphology, a small volume of the formalin fixed sperm was placed onto a clean glass slide and allowed to air dry. Sperm cells were then examined at 160x magnification using a Leica DM6000 B digital light microscope (Leica Microsystems, Switzerland) and digital images were captured with a Leica DFC420 camera (Leica Microsystems, Switzerland). For each male, 10 morphologically normal sperm were examined and the following measurements recorded: head, midpiece and tail length (to the nearest ±0.1 μm). From these measurements the following two additional metrics were also calculated: flagellum length (i.e. midpiece + tail length) and total sperm length (i.e. head + midpiece + tail length). For each individual male we calculated an average value for each sperm trait. We also calculated values for within-male (CV_{wm}) and between-male (CV_{bm}) coefficient of variation in total sperm length. Finally, because the CV_{bm} measure has been documented to

be underestimated in small population samples (Laskemoen *et al.*, 2007), values of CV_{bm} were adjusted according to sample size using the following formula: $CV_{bm} + (1/(4n))$ (Sokal & Rohlf, 1995).

4 Summary of papers

4.1 Paper I: Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*.

In this paper, we investigated the deep divergence found in mtDNA in the common redstart. Our primary goal was to examine this variation with reference to the process of speciation (i.e. understanding whether such divergence represents reversed speciation or early speciation) and processes such as hybridization and cryptic speciation. As such, this paper had the following two main aims. First, we quantified the geographical distribution of the two mtDNA haplogroups found in the common redstart (Johnsen *et al.*, 2010) and examined the degree of sympatry and interbreeding observed across the species' breeding range. Second, we considered five hypotheses that could explain how such deep mtDNA divergence (~ 5% K2P distance) may have originated, and tested a range of predictions underlying each of these hypotheses: (1) Amplifying non-functional copies of mtDNA (numts) hypothesis – support for this hypothesis would come from the presence of stop codons and multiple double peaks in the mitochondrial sequences implying the presence of numts. (2) Cryptic species hypothesis – support for this hypothesis would come from the occurrence of assortative mating or differences in the sperm morphology between the haplogroups. More specifically, the occurrence of these traits would suggest the presence of reproductive barriers and therefore the existence of cryptic species. Additionally, divergence in nuclear DNA associated with a degree of divergence in mtDNA would also suggest the presence of cryptic species. (3) Hybridization hypothesis – the hypothesis of hybridization between taxa would be supported if one or more of the haplogroups (based on COI) cluster together with another extant *Phoenicurus* species. (4) Geographic isolation hypothesis – evidence of structure in the geographical distribution of the two haplogroups, along with different mismatch distributions resulting from different demographic histories would support the hypothesis that haplotypes have been geographically isolated in the past. Furthermore, a lack of reproductive barriers combined with little or no structure in the

nuclear sequences (with respect to the mtDNA lineages) despite high nuclear variation would suggest speciation in reverse. Evidence for reproductive isolating mechanisms combined with a pattern of divergence in the nuclear data as a result of the two lineages being effectively separated would, on the other hand, suggest that the haplotypes are in the process of early speciation. Finally, (5) Co-existence / panmictic population hypothesis – support for the hypothesis that the haplotypes belong to a single panmictic population would come from a lack of geographic population structure and the absence of reproductive barriers, as well as similar mismatch distributions and no divergence in the nuclear introns.

These hypotheses were tested by combining sequence data from two mtDNA regions (control region and COI) and two nuclear Z-linked introns (BRM-15 and ALDOB-6), with data on assortative mating and variation in sperm morphology for the two haplogroups. Because we found no evidence for mitochondrial pseudogenes, lineage-specific assortative mating, or variation in sperm morphology, as well as no evidence of hybridization with an extant *Phoenicurus* species (based on a short fragment of the COI), we excluded the first three hypotheses. However, introgression from an extinct congeneric cannot be excluded. Next, while mitochondrial sequencing revealed two distinct haplogrops, sequenced nuclear introns failed to show such distinct groupings. Finally, similar mismatch distributions were found for the two mitochondrial haplogroups, suggesting that the two mitochondrial lineages have undergone similar recent demographic changes. From these results, we concluded that the deep, sympatric mtDNA lineages found in the common redstart did not represent cryptic species, nor were they likely to result from introgression from extant congenerics. Rather, the data suggested that haplotype divergence either evolved in isolated refugia with subsequent secondary contact or represented ancestral lineages that coexisted in one panmictic population, or some combination of these two scenarios.

4.2 Paper II: Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*, Lepidoptera, Geometridae).

Here, an ecologically important moth species, the autumnal moth, was used to investigate high mtDNA divergence. Importantly, examination of high mtDNA divergence in both a moth and an avian species allowed us to investigate whether the processes underlying this pattern are consistent across highly divergent taxa. Additionally, we aimed to understand if

and how infection by an endosymbiont (i.e. *Wolbachia*) influences the degree of mtDNA variation found within this species.

To achieve our goal, we examined the high mtDNA variation found within the autumnal moth, described the degree of sympatry among haplogroups within Norwegian populations of this species and compared variation in mtDNA to variation in nuclear loci. More specifically, we investigated four possible explanations for the occurrence of high intraspecific mtDNA variation: (1) Assuming sufficient time for divergence, congruence between divergence in mtDNA and nuclear DNA sequence data would suggest the occurrence of cryptic species. (2) Higher differentiation in mtDNA compared to nuclear DNA, based on the relatively high evolutionary rate of mtDNA (Avise *et al.*, 1988), would suggest that variation is due to isolation and possibly repeated secondary contact occurring long ago. Moreover, assuming sufficient time since range expansion and secondary contact, the degree of mtDNA and nuclear DNA structure would reflect the demographic history and original geographic distribution of the lineages. (3) Higher differentiation in mtDNA compared to nuclear DNA would also be consistent with mtDNA introgression by hybridization. Furthermore, the occurrence of overlapping haplotypes with closely related species (e.g. *E. dilutata* and/or *E.christyi*) would indicate that introgression occurred relatively recently in evolutionary time. Finally, (4) an association between *Wolbachia* infection status and haplogroups and incongruence between mtDNA and nuclear DNA would suggest that *Wolbachia* infections have influenced mtDNA variation in this species. Consequently, we screened samples for *Wolbachia* to determine whether *Wolbachia* has influenced patterns of mitochondrial diversity in the autumnal moth.

All individuals were sequenced using a mitochondrial marker (COI), and a subset of individuals was sequenced using nuclear markers (ITS2 and wingless) in order to resolve the discrepancy found between the mtDNA divergence and present species-level taxonomy. A total of five sub-clades were found in the COI region within the autumnal moth complex (divergence 1.5-4.1%). The majority of these sub-clades were sympatric and showed little geographic variation. In the nuclear markers little variation was found, and there was no indication of more than one species present. When screening for *Wolbachia* infections, 12 % of the samples tested positive, and two *Wolbachia* strains were associated with different mtDNA sub-clades which may indicate indirect selection/selective sweeps on these haplotypes. Thus we concluded that the most likely explanation for the high mitochondrial variation is that current populations consist of separate lineages that once

evolved in allopatry, without evolving reproductive barriers, and that *Wolbachia* infections may contribute to maintaining this variation in sympatric populations today.

4.3 Paper III: Rapid sperm evolution in the bluethroat (*Luscinia svecica*) subspecies complex.

In this paper, we aimed to investigate whether evolution in a subspecies complex could be detected using variation in sperm morphology, and to relate this variation to the early stages of speciation, using five (*Luscinia svecica azuricollis, L. s. cyanecula, L. s. namnetum, L. s. svecica and L. s. volgae*) subspecies of bluethroats. More specifically, this paper had two main aims: 1) To test if there is a relationship between genetic divergence and sperm divergence within these five subspecies (across seven study populations), and 2) To compare differences in sperm evolution between bluethroat and other species with known sperm divergence.

Our results showed that these subspecies exhibit small genetic divergences in neutral markers, and analyses of mtDNA suggests that this subspecies divergence is very recent (maximal genetic distance, 0.7% = 350 000 years ago). We found significant variation in total sperm length and in the length of some sperm components (i.e. head and midpiece) among the subspecies, and a significant correlation between genetic divergence and divergence in total sperm length among the subspecies. The degree of divergence in sperm morphology found between the bluethroat subspecies was considerably higher than those observed between both sister species and other populations/subspecies groups. Taken together, we suggest that high divergence in sperm morphology, combined with low genetic divergence, indicates rapid evolution of sperm traits in this system. Finally, we considered the relative role of selection (e.g. sperm competition) and genetic drift in sperm divergence, and suggest that sperm divergence may play an important role in the early stages of the speciation process.

4.4 Paper IV: Increased divergence but reduced variation on the Z chromosome relative to autosomes in *Ficedula* flycatchers: differential introgression or the faster-Z effect?

We conducted this study in order to investigate whether higher variation on Z chromosomes relative to autosomes, a pattern first described in the pied (*Ficedula hypoleuca*) and collared (*F. albicollis*) flycatchers (Borge *et al.*, 2005b), occurs in the black-and-white *Ficedula* species complex more generally (i.e. in all four species). More

specifically, this paper aimed to test two hypotheses regarding differences in variation between Z-linked and autosomal loci using the *Ficedula* flycatcher species. These two hypotheses were: (1) The faster-Z hypothesis (Charlesworth *et al.*, 1987), and (2) The differential introgression hypothesis (Carling *et al.*, 2010; Storchova *et al.*, 2010; Backström & Väli, 2011).

The faster-Z hypothesis states that faster adaptive evolution on the Z chromosome is expected because (partially) recessive beneficial mutations are not masked by dominance in the heterogametic sex. Likewise, (partially) recessive deleterious mutations would be more effectively purged on the Z compared to autosomes due to hemizygous exposure. Furthermore, associated selective sweeps on the Z chromosome are expected to further contribute to reductions in intraspecific polymorphism (Charlesworth *et al.*, 1987; Borge *et al.*, 2005b). Genetic drift may also contribute to a faster-Z effect because the lower effective population size of the Z chromosome would be associated with increased rates of genetic drift and thus an increased fixation rate of mildly deleterious mutations (Charlesworth *et al.*, 1987; Mank *et al.*, 2010).

The differential introgression hypothesis states that the accumulation of incompatibilities on the Z chromosome may reduce the rate of introgression of Z-linked genes compared to autosomal genes, resulting in the same pattern as predicted by the faster-Z hypothesis (Carling *et al.*, 2010; Storchova *et al.*, 2010; Backström & Väli, 2011). Thus, the two hypotheses are not mutually exclusive and determining the exact reason for such a pattern is difficult. For instance, a faster-Z effect may speed up divergence and hence contribute to the accumulation of sex-linked incompatibilities that would reduce Z-linked introgression (e.g. Elgvin *et al.*, 2011). Nevertheless, this paper employed recently developed isolation with migration (IMa) models in order to ascertain the most likely underlying cause for the patterns observed in the flycatchers.

We found that the *Ficedula* flycatchers show greater genetic divergence on the Z chromosome than the autosomes, and that the ratios of intraspecific polymorphism at Z-linked vs. autosomal markers were below the neutral expectation of 0.75%. Additionally, using isolation with migration (IMa) models we estimated gene flow among the four closely related flycatcher species. Our results suggest that the patterns found here can best be explained by the faster-Z hypothesis, since the estimated long-term gene flow parameters were close to zero in all comparisons.

5 Discussion

This thesis presents three main findings. First, high mitochondrial divergence was found in the sympatric populations of the common redstart and the autumnal moth. This divergence was not synonymous with the presence of pseudogenes, early speciation or cryptic speciation in either study system. However, in both the common redstart and the autumnal moth, we suggested that this divergence may reflect populations that have evolved in isolated refugia with secondary contact occurring at a later point in time, without evolving reproductive barriers (speciation in reverse), and that the divergence in the autumnal moth has been maintained by *Wolbachia* infections. Second, high divergence was found in both primary and secondary sexual characters in bluethroat subspecies. This finding, combined with the low divergence in mitochondrial DNA, prompted us to suggest that these sexual characters (i.e. sperm and throat coloration) have undergone rapid evolutionary change and that the subspecies are in an early stage of the speciation process. Third and finally, we found that high variation combined with reduced divergence on the Z chromosome compared to the autosomes in the young *Ficedula* flycatcher species complex was best be explained by the faster-Z hypothesis.

5.1 Speciation in reverse

The most common pattern found in phylogeography is the splitting of lineages into new species, while remerging of divergent lineages (i.e. reversed speciation) is reported relatively infrequently (Seehausen *et al.*, 1997; Turner, 2002; Taylor *et al.*, 2006; Webb *et al.*, 2011). Four main explanations for the high mitochondrial divergence found in sympatric populations are common to both the redstart and the autumnal moth studies: 1) The occurrence of pseudogenes, 2) Cryptic speciation, 3) Hybridization, and 4) Speciation in reverse. For both paper I and II, we screened sequences for double peaks and stop codons, but found evidence of neither. Consequently, the pseudogene hypothesis was excluded in both studies. Similarly, we searched for evidence of cryptic speciation by investigating reproductive barriers and assortative barriers, but found none. To test for hybridization, we compared the mitochondrial sequences from each species with closely related species. As before, we found no evidence to support either scenario and thus both cryptic speciation and hybridization were ruled out. Instead, we found the most likely scenario in both study systems to be one where high divergence was either established in

isolated refugia, without the evolution of reproductive barriers, and maintained upon secondary contact in sympatry later (speciation in reverse), or that ancestral lineages coexisted in one panmictic population. Finally, a combination of these two scenarios may also explain our findings. In addition, we found evidence that mitochondrial variation was associated with *Wolbachia* infections in the autumnal moth. Endoparasite infections have been suggested to maintain high mtDNA divergence in other insects (e.g. Eurasian two-spot ladybirds, *Adalia bipunctata*; *(Schulenburg et al., 2002)*, supporting our suggestion that different *Wolbachia* strains contribute to the maintenance of the lineages found in the autumnal moth, and that such lineages are maintained due to indirect selection on different haplotypes.

5.2 Early speciation/rapid evolution of phenotypic traits

The process of lineage splitting is a well-known phenomenon which may occur in several ways, e.g. through subdivision by the appearance of a barrier or because of a rare dispersal event across a barrier (Price, 2008). Based on mitochondrial markers, the five bluethroat subspecies studied in this thesis show evidence of recent divergence (maximum genetic distance 0.7% = 350 000 years ago, paper III). In contrast to this low genetic divergence, there are considerable differences in both sperm morphology (paper III) and male throat coloration (e.g. Johnsen *et al.*, 2006). Few studies of sperm morphology have been conducted on subspecies complexes. However, in a study of the barn swallow (*Hirundo rustica*), Laskemoen *et al.* (in press), found differences in sperm morphology between subspecies, and argued that these differences might result from either genetic drift or selection (or a combination of the two factors). Moreover, the authors suggested that sperm morphology could provide an indication of genetic distance between species and between lineages within species (Laskemoen *et al*., in press). In paper III, we found a positive and significant relationship between sperm morphology and mitochondrial DNA, suggesting that drift may explain some of the variance in sperm morphology found within the bluethroat subspecies. Nonetheless selection by sperm competition cannot be ruled out as an influential factor in the evolution of sperm traits in this subspecies complex. Taken together, the low genetic divergence found between the bluethroat subspecies combined with the more rapid divergence in both sperm morphology and throat coloration suggest that these subspecies are in the early stages of speciation.

5.3 Recent speciation, Z chromosome evolution and the faster-Z hypothesis

The black-and-white *Ficedula* flycatchers are four well-studied species that have undergone recent speciation. The collared and pied flycatchers in particular have been studied intensively with regards to speciation ecology, behavior and genetics (Qvarnström *et al.*, 2010; Sætre & Sæther, 2010). Despite repeated episodes of contact in the past, the mitochondrial genomes of the four *Ficedula* flycatchers are clearly divergent (Sætre *et al.*, 2001; Sætre *et al.*, 2003; Borge *et al.*, 2005b), suggesting a history of restricted gene flow between allopatric populations. Two of the flycatchers (collared and pied), however, do have overlapping distributions in both central Europe and on the islands of Öland and Gotland in Sweden, where they are known to interbreed. Between 2% and 7% of breeding flycatchers are hybrids according to estimates from different mixed-species populations on these islands (Alatalo *et al.*, 1990; Sætre *et al.*, 1999; Veen *et al.*, 2001). In these populations, females appear to be sterile, while males show signs of reduced reproductive fitness. Low introgression on the Z chromosome relative to the autosomes and limited intraspecific recombination on the Z chromosome have been shown in flycatchers in hybrid zones (Sætre *et al.*, 2003; Borge *et al.*, 2005a). In paper IV, we show that the pattern of fewer shared polymorphisms and more fixed differences on Z-linked genes holds for all four of the *Ficedula* flycatchers. Such a pattern of increased divergence and reduced polymorphism has been shown in other closely related bird species (Berlin & Ellegren, 2004; Storchova *et al.*, 2010; Backström & Väli, 2011; Elgvin *et al.*, 2011), and may be explained by faster adaptive divergence on the Z chromosome (the faster-Z hypothesis) or reduced introgression on the Z chromosome due to accumulation of sex-linked incompatibilities (the differential introgression hypothesis). In paper IV, we suggested (based on IMa analyses) that this pattern in the flycatcher is best explained by the faster-Z hypothesis. Earlier studies on the pied and collared flycatcher show that Z-linked genes are related to reproductive success in these species, since both traits involved in premating isolation (Sæther *et al.*, 2007) and post-zygotic barriers (Sætre *et al.*, 2003) appears to be associated with Z-linked genes in flycatchers. The pattern found here might reflect past episodes of secondary contact in hybrid zones, where the Z chromosome and autosomes have experienced different levels of introgression, and the substantial divergence on the Z chromosome is likely to have implications for speciation processes in the four *Ficedula* flycatchers.

5.4 DNA barcoding

In this thesis, the barcoding region (COI) was used in all four study groups (paper I, II, III, and figure 3 for flycatchers), the results of which have important implications for DNA barcoding and the detection of cryptic species. DNA barcoding was first proposed by Hebert (2003) as a method using a short section of a standardized region of the mitochondrial genome (COI) to identify and discover species. Subsequent efforts have led to the development of a DNA barcode library (Ratnasingham & Hebert, 2007), which is a repository for sequences from different taxa and can be used to identify material from unknown sources (e.g. partial samples, organisms that are difficult to distinguish using morphological traits). Moreover, the barcode library is thought to allow for the easy determination of unidentified samples and the recognition of new species. For most animal taxa, DNA barcoding makes use of the COI region (Hebert *et al.*, 2003), while two regions in the chloroplast (matK and rbcL) have been suggested as the barcoding region of choice for land plants (Hollingsworth *et al.*, 2009). In this thesis, DNA barcoding worked well for the *Ficedula* flycatchers: all four species showed monophyletic clades with a divergence of ~ 2% (Figure 3, *Ficedula parva* as outgroup), making them easy to distinguish and COI a reliable marker of species identification. In contrast, in both the redstart and the autumnal moth, we found high levels of intraspecific divergence, which would suggest these groups comprise two (redstart) and five (autumnal moth) distinct species (with COI divergence ~ 5% and between 1.5-4.1%, respectively) according to Hebert *et al.*'s (2003) criterion. However, as there is no indication of either of these taxa comprising more than one species, DNA barcoding does not in this instance provide a reliable classification system. Finally, the five bluethroat subspecies examined in this thesis showed very little divergence in COI (0.3-0.5%) making them impossible to discriminate using COI alone.

Several studies have criticized DNA barcoding for relying too much on one gene and for using a threshold value for identification of species (Moritz & Cicero, 2004; Meyer & Paulay, 2005; Hickerson *et al.*, 2006). The results of our studies also suggest that a threshold value may be difficult to use for effective species delineation, and consequently we suggest that care should be taken when analyzing new species using this single gene system. Nonetheless, COI is a valuable tool for identifying most species, but in instances where more than one haplogroup is found within a species, the use of COI should be supported with evidence from morphological traits, additional genetic sampling or a combination of both of these factors (DeSalle *et al.*, 2005; Damm *et al.*, 2010; Dupuis *et al.*, 2012). This thesis also demonstrates that COI may not be a suitable approach to

subspecies discrimination. For example, for the bluethroat subspecies investigated, COI identifies the correct species (i.e. *L. svecica*), and to some degree can identify sequences from the subspecies *L. s. svecica* or *L. s. azuricollis*, but based on COI alone it is currently impossible to identify subspecies with certainty. The same applies for the common redstart, for which DNA barcoding is unable to identify the two described subspecies (*P. p. phoenicurus* and *P. p. samamisicus*). Nevertheless, species level identification is correct using COI data.

5.5 Genetic divergence and the process of speciation

All species included in this thesis are considered well-recognized species. The discrepancy between the genetic distances found in the mitochondrial genes within these species (~ 5% for the redstarts, 1.5-4.1% for the autumnal moth), between the flycatchers species (2.1-2.9%) and between subspecies (0.7% maximum genetic distance within the bluethroat subspecies) may be (at least partially) explained by them being in different stages of the speciation process. The redstart and the autumnal moth both show high mitochondrial divergence in sympatric populations. This divergence has been suggested in the current studies (paper I and II) to be a result of speciation in reverse, which in the case of the autumnal moth may be slowed by *Wolbachia* infections (i.e. *Wolbachia* appears to maintain current levels of intraspecific variation). Speciation in reverse has only rarely been demonstrated in avian taxa, one exception being the common raven (*Corvus corax*) for which isolation followed by speciation in reverse has been suggested as the major explanation for high sympatric divergence in the mitochondrial genome (Webb *et al.*, 2011). Thus this thesis adds another valuable example of this phenomenon to the published literature. In the bluethroat subspecies complex (paper III) we found evidence of early speciation; sperm morphology appears to have undergone rapid evolutionary changes despite low genetic divergence in molecular traits.

The ecology of a species has been suggested to contribute to the process of speciation, at least under conditions of allopatry. Specifically, ecological specialist species are suggested to be less likely to remerge upon secondary contact because the probability of retaining these evolved specializations is high (Price, 2008; Schluter, 2009). In contrast, species that have evolved to be ecological generalists are less likely to exhibit unique adaptations (due to their wide ecological tolerance) and, as such, are more likely to remerge after secondary contact has been established (Webb *et al.*, 2011). Consequently, ecology may provide a partial explanation for the high mitochondrial divergence observed

in the redstart and autumnal moth, and the low divergence observed in the bluethroats (which may be in an early speciation stage). For example, the redstart appears to exhibit preference for a wider range of ecological habitats across the entire breeding range; they require fairly open wooded or parkland areas with access to dry nest-holes and sparse undergrowth (Cramp, 1988b), suggesting it can be considered an ecological generalist. In contrast, at least some of the recognized bluethroat subspecies exhibit preference for breeding areas that vary in ecology. The subspecies *L. s. svecica* has been shown to prefer wooded tundra with marshy glades, while *L. s. cyanecula* prefers lowland reed beds and *L. s. azuricollis* prefers dry stony slopes up to 2000m, covered with Spanish broom (Cramp, 1988a). Finally, the *L. s. namnetum* exhibits a preference for salt marsh areas surrounded by patches of bushes and small trees (Allano *et al.*, 1994). Thus the bluethroat may represent a more specialized group of subspecies, which has in turn lead to lower probability of collapse between populations.

The low divergences found in neutral nuclear markers (paper I, II and III) may be explained by these markers being less likely to show divergence, due to both recombination and lower rates of evolutionary change (relative to mtDNA), which will more easily allow for mixing when the two populations come into secondary contact, provided of course that these populations have not evolved effective reproductive barriers (Webb *et al.* 2011).

5.6 Speciation and sexual selection

A paradox found in this thesis is that the species showing the lowest levels of mitochondrial divergence are those that appear to have progressed the furthest in terms of the speciation process. Thus, the following major question arises from this thesis: What might explain the association between low mitochondrial divergence and progress towards speciation? I suggest that this paradox may be (at least partially) explained by variation in sexual selection pressures faced by species, which are related to differences in life history. For example, the common redstart is a sexually dimorphic species consisting of two subspecies that currently show no genetic variation (paper I). This species is primarily characterized as socially monogamous (Cramp, 1988b) and appears to experience low level of sperm competition (Kleven *et al.*, 2007). Moreover, there is no evidence of either haplotype-associated assortative mating or divergence in sperm morphology in this species. Thus I suggest the common redstart experiences a relatively low intensity of sexual selection. Similarly, the autumnal moth exhibits little morphological divergence between

the two sexes, though a slight difference in abdominal size does occur (Hausmann & Viidalepp, 2012). However, these differences are minor and are difficult to observe by visual inspection alone. The level of sperm competition and sexual selection is unknown for the autumnal moth, but a lack of divergence in the genitalia of the haplogroups (Kvie, 2010) suggests such selective pressures are likely to be low. In contrast, the bluethroat exhibits strong sexual dichromatism (generally drab plumed females and colorful males; Johnsen *et al.* 2006) and differences in both plumage coloration and sperm morphology are found between the subspecies. These differences may imply that the populations have experienced divergent sexual selection pressures in allopatry. In addition, the bluethroat exhibits high levels of sperm competition (Questiau *et al.*, 1999; Johnsen & Lifjeld, 2003; paper III), and polygyny is relatively common (A. Johnsen & J. T. Lifjeld unpublished). Taken together, these traits suggest that the bluethroat subspecies complex experiences intense sexual selection, and this intense selection may explain why this species has come further in the speciation process than both the common redstart and autumnal moth. Finally, the black-and-white flycatchers consist of four species that exhibit some differences in morphological characters (predominately male plumage traits), and both premating and postmating barriers have been identified in these species (reviewed in Qvarnström *et al.*, 2010; Sætre & Sæther, 2010). Moreover, the level of sperm competition has been shown to be moderate in the collared (Sheldon & Ellegren, 1999; Krist *et al.*, 2005) and pied flycatcher (Lifjeld *et al.*, 1991; Rätti *et al.*, 1995), though polygyny is common in the latter species (Slagsvold & Lifjeld, 1986), suggesting that this species faces a moderate level of sexual selection. Thus, across these four studies, it appears that species experiencing more intense sexual selection exhibit the greatest degree of progress in the stages of speciation, suggesting that sexual selection has influenced the rate of evolutionary diversifications among clades studied in this thesis

As suggested above, sexual selection may explain why the species with lowest COI divergence has come furthest in the speciation process. Darwin (1871) was the first to suggest that sexual selection may contribute to increasing the diversity of a clade, and this theory has gained popularity during recent years (Andersson, 1994; Barraclough *et al.*, 1995; Panhuis *et al.*, 2001). More specifically, assortative mating and reproductive barriers are suggested to evolve when there is divergence in male traits and the expression of female preferences for these traits between populations of a single species and if a genetic correlation is established between male traits and female preferences (Andersson, 1994). Several studies have investigated sexual dimorphism in birds in relation to speciation, and

found a positive relationship between the number of species within a clade and the degree of dimorphism (Barraclough *et al.*, 1995; Mitra *et al.*, 1996; Møller & Cuervo, 1998; Seddon *et al.*, 2008; but see Morrow & Pitcher, 2003; Morrow *et al.*, 2003). A recent study of the barn swallow demonstrates how such patterns of phenotypic variation in a species complex could be driven by differential sexual selection pressures in different populations; in this species tail length and ventral coloration varies among populations, as does female preference for these traits (Vortman *et al.*, 2011). This may also be the case for the bluethroat. Specifically, males in the different bluethroat subspecies may be experiencing differential sexual selection pressures for both primary and secondary sexual characters (e.g. sperm morphology and throat coloration), making this subspecies complex more morphologically diverse relative to the redstarts, in which sexual selection is less intense. This idea is consistent with the hypothesis that the bluethroats are at an early stage of the speciation process and that sexual selection drives evolutionary diversification and creates a pattern of high morphological differentiation and low genetic divergence, while the redstart show signs of speciation in reverse.

6 Conclusions

The species investigated in this thesis appear to be in different stages of the speciation process. The two species with highest mitochondrial divergence both show signs of speciation in reverse, the subspecies complex with lowest genetic divergence show signs of early speciation, while the four *Ficedula* flycatchers probably are a result of recent speciation. I have suggested that both differences in sexual selection pressures and ecological adaptations might explain these differences. In addition, the study species show contrasting patterns of sperm evolution; specifically the high divergence observed in total sperm length between the bluethroat subspecies suggests that taxa in this clade have experienced rapid evolutionary change in sperm traits. Finally, the contrasting pattern found between the Z chromosome and autosomes in the *Ficedula* flycatchers, the increased divergence and reduced variation on Z, is best explained by the faster-Z hypothesis. In conclusion, by combining knowledge of morphology, genetics and ecology, I have been able to develop a more comprehensive understanding of the early stages of the speciation process. Future work on speciation would benefit from using such an integrated approach.

7 Future prospects

In this thesis, I have used traditional sequencing methods to show that variation in mitochondrial divergence can be connected to different stages of the speciation process. In all four study groups, I suggest that future research efforts should use next generation sequencing to obtain a deeper genetic coverage. Such efforts would provide valuable information on nuclear DNA evolution and help explain discrepancies in divergence between nuclear and mitochondrial regions.

For the autumnal moth it would be interesting to further investigate the role of *Wolbachia* in the maintenance of intraspecific diversity in this species. In addition, the acquisition of samples from a wider range of the moth's distribution area would offer a better overview of the genetic structure of *E. autumnata*. Moreover, sequencing of nuclear markers that evolve faster than the loci studied here and testing for further ecological and morphological differences between variants would provide more information regarding the role of geography and ecology in maintaining the observed genetic variation in *E. autumnata*.

For the bluethroats, it would be interesting to use next generation sequencing in order to obtain a well-resolved phylogeny for the subspecies. Such a phylogeny would be important in order to get a better understanding of the rapid sperm evolution in this species, and to investigate the directionality of evolutionary change in sperm traits. Another interesting question that remains unanswered is whether allopatric divergence in sperm morphology actually functions as a barrier to gene flow between the different subspecies. This could be tested for by performing *in vitro* experiments, mixing sperm and female fluids between subspecies, and testing for effects on sperm motility or (though technically challenging) through the use of artificial insemination techniques to cross-inseminate females with sperm from males belonging to different subspecies.

Additionally, further investigation of sexual selection for male throat ornamentation is warranted, including identifying the underlying genetic causes of color variation. Importantly, the central spot of the throat ornament has been shown to be a putative subspecies discrimination cue in this species (Johnsen *et al.* 2006). Thus, I recommend future experimental work should attempt to manipulate male throat coloration. Specifically, one could, for example manipulate throat coloration in a red-spotted populations to resemble males of the central-European white-spotted subspecies and vice versa, and test

whether such manipulations impacts male mating success (e.g. pairing success, within pair paternity and extrapair paternity success). In addition, it would be interesting to analyze microstructure and pigment content of the ornamental feathers of the most distinct subspecies and identify genetic structure of this variation to examine the possible sequence of evolutionary change in this character.

8 References

Aarvik, L., Hansen, L. O. & Kononenko, V. 2009. *Norges sommerfugler, Håndbok over Norges dagsommerfugler og nattsvermere*. Norsk entomologisk forening, Naturhistorisk museum, Oslo.

Alatalo, R. V., Eriksson, D., Gustafsson, L. & Lundberg, A. 1990. Hybridization between pied and collared flycatchers- sexual selection ans speciation theory. *Journal of Evolutionary Biology* 3: 375-389.

Allano, L., Bonnet, P., Constant, P. & Eybert, M. C. 1994. Structure de l'habitat et desité de population chez la Gorgebleue à miroir *Luscinia-svecica namnetum* Mayaud. *Revue D Ecologie-La Terre Et La Vie* 49: 21-33.

Andersson, M. 1994. *Sexual selection*. Princeton University Press, Princeton, New Jersey.

Avise, J. C. 2000. *Phylogeography*. Harvard University press, Cambridge, Massachusetts London, England.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E. *et al.* 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between populaiton-genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522.

Avise, J. C., Ball, R. M. & Arnold, J. 1988. Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molecular Biology and Evolution* 5: 331-344.

Backström, N., Lindell, J., Zhang, Y., Palkopoulou, E., Qvarnstrom, A., Sætre, G. P. *et al.* 2010. A high-density scan of the Z chromosome in *Ficedula* flycatchers reveals candidate loci for diversifying selection. *Evolution* 64: 3461-3475.

Backström, N. & Väli, U. 2011. Sex- and species-biased gene flow in a spotted eagle hybrid zone. *Bmc Evolutionary Biology* 11.

Ballard, J. W. O. & Whitlock, M. C. 2004. The incomplete natural history of mitochondria. *Molecular Ecology* 13: 729-744.

Barraclough, T. G., Harvey, P. H. & Nee, S. 1995. Sexual selection and taxonomic diversity in passerine birds. *Proceedings of the Royal Society B-Biological Sciences* 259: 211-215.

Bazin, E., Glémin, S. & Galtier, N. 2006. Population Size Does Not Influence Mitochondrial Genetic Diversity in Animals. *Science* 312: 570-572.

Berlin, S. & Ellegren, H. 2004. Chicken W: A genetically uniform chromosome in a highly variable genome. *Proceedings of the National Academy of Sciences of the United States of America* 101: 15967-15969.

Bermingham, E. & Moritz, C. 1998. Comparative phylogeography: concepts and applications. *Molecular Ecology* 7: 367-369.

Borge, T., Lindroos, K., Nadvornik, P., Syvänen, A. C. & Sætre, G. P. 2005a. Amount of introgression in flycatcher hybrid zones reflects regional differences in pre and post-zygotic barriers to gene exchange. *Journal of Evolutionary Biology* 18: 1416-1424.

Borge, T., Webster, M. T., Andersson, G. & Sætre, G. P. 2005b. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* 171: 1861-1873.

Bush, G. L. 1994. Sympatric speciation in animals- new wine in old bottles. *Trends in Ecology & Evolution* 9: 285-288.

Carling, M. D., Lovette, I. J. & Brumfield, R. T. 2010. Historical divergence and gene flom: coalescent analyses of mitochondrial, autosomal and sex-linked loci in *passerina* buntings. *Evolution* 64: 1762-1772.

Charlesworth, B., Coyne, J. A. & Barton, N. H. 1987. The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist* 130: 113-135.

Coyne, J. A. & Orr, A. H. 2004. *Speciation*. Sinauser Associates, Inc., Sunderland.

Cramp, S. (1988a) Bluethroat. In: *The Birds of the Western Palearcitc. Handbook of Birds in Europe, the Middle East and Africa*, (Cramp, S., ed.). pp. 645-661. Oxford University Press, Oxford.

Cramp, S. (1988b) Redstart. In: *The Birds of the Western Palearcitc. Handbook of Birds in Europe, the Middle East and Africa*, (Cramp, S., ed.). pp. 695-708. Oxford University Pree, Oxford.

Damm, S., Schierwater, B. & Hadrys, H. 2010. An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology. *Molecular Ecology* 19: 3881-3893.

Darwin, C. 1871. *The decent of man, and selection in relation to sex*. Murray, London.

del Hoyo, J., Elliot, A. & Christie, D. A. 2005. *Handbook of the birds of the world*. Lynx Edicions, Barcelona.

DeSalle, R., Egan, M. G. & Siddall, M. 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society B-Biological Sciences* 360: 1905-1916.

Dupuis, J. R., Roe, A. D. & Sperling, F. A. H. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology* 21: 4422-4436.

Edwards, S. V., Kingan, S. B., Calkins, J. D., Balakrishnan, C. N., Jennings, W. B., Swanson, W. J. *et al.* 2005. Speciation in birds: Genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences of the United States of America* 102: 6550-6557.

Elgvin, T. O., Hermansen, J. S., Fijarczyk, A., Bonnet, T., Borge, T., Sæther, S. A. *et al.* 2011. Hybrid speciation in sparrows II: a role for sex chromosomes? *Molecular Ecology* 20: 3823-3837.

Ellegren, H. 2009. The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* 25: 278-284.

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T. *et al.* 2012. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* 491: 756-760.

Endler, J. A. 1977. *Geographic variation, speciation, and clines*. Princeton University Press, Princeton, N.J.

Gilbert, W. 1978. Why genes in pieces? *Nature* 271: 501-501.

Hausmann, A. & Viidalepp, J. (2012) Larentiinae I. In: *The Geometrid Moths of Europe*, Vol. 3 (Hausmann, A., ed.). pp. 357-359. Apollo Books, Denmark.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B-Biological Sciences* 270: 313-321.

Hewitt, G. M. 2001. Speciation, hybrid zones and phylogeography — or seeing genes in space and time. *Molecular Ecology* 10: 537-549.

Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359: 183-195.

Hickerson, M. J., Meyer, C. P. & Moritz, C. 2006. DNA barcoding will often fail to discover new animal species over broad parameter space. *Systematic Biology* 55: 729-739.

Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., van der Bank, M. *et al.* 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 12794-12797.

Johnsen, A., Andersson, S., Fernandez, J. G., Kempenaers, B., Pavel, V., Questiau, S. *et al.* 2006. Molecular and phenotypic divergence in the bluethroat (*Luscinia svecica*) subspecies complex. *Molecular Ecology* 15: 4033-4047.

Johnsen, A. & Lifjeld, J. T. 2003. Ecological constraints on extra-pair paternity in the bluethroat. *Oecologia* 136: 476-483.

Johnsen, A., Rindal, E., Ericson, P., Zuccon, D., Kerr, K., Stoeckle, M. *et al.* 2010. DNA barcoding of Scandinavian birds reveals divergent lineages in trans-Atlantic species. *Journal of Ornithology* 151: 565-578.

Kleven, O., Laskemoen, T., Fossøy, F., Robertson, R. J. & Lifjeld, J. T. 2008. Intraspecific variation in sperm length is negatively related to sperm competition in passerine birds. *Evolution* 62: 494-499.

Kleven, O., Øigarden, T., Foyn, B. E., Moksnes, A., Røskaft, E., Rudolfsen, G. *et al.* 2007. Low frequency of extrapair paternity in the common redstart (*Phoenicurus phoenicurus*). *Journal of Ornithology* 148: 373-378.

Krist, M., Nadvornik, P., Uvirova, L. & Bures, S. 2005. Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology* 59: 6-11.

Krokene, C., Anthonisen, K., Lifjeld, J. T. & Amundsen, T. 1996. Paternity and paternity assurance behaviour in the bluethroat, *Luscinia s svecica*. *Animal Behaviour* 52: 405-417.

Kvie, K. S. (2010) Epirrita autumnata (Lepidoptera, Geometridae): one diverse species or a complex of several cryptic species? Master of Science thesis. Natural History Museum, UiO.

Kvist, L., Martens, J., Nazarenko, A. A. & Orell, M. 2003. Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution* 20: 243-247.

Laskemoen, T., Albrecht, T., Bonisoli-Alquati, A., Cepak, J., De Lope, F., Hermosell, I. G. *et al.* in press. Variation in sperm morphometry and sperm competition among barn swallow (*Hirundo rustica*) populations. *Behavioral Ecology and Sociobiology*: DOI: 10.1007/s00265-012-1450-0.

Laskemoen, T., Kleven, O., Fossøy, F. & Lifjeld, J. T. 2007. Intraspecific variation in sperm length in two passerine species, the bluethroat *Luscinia svecica* and the willow warbler *Phylloscopus trochilus*. . *Ornis Fennica* 84: 131-139.

Lifjeld, J. T., Slagsvold, T. & Lampe, H. M. 1991. Low-frequency of extra-pair paternity in pied flycatchers revealed by DNA fingerprinting. *Behavioral Ecology and Sociobiology* 29: 95-101.

Lundberg, A. & Alatalo, R. V. 1992. *The Pied flycatcher*. Poyser, London.

Mank, J. E., Nam, K. & Ellegren, H. 2010. Faster-Z evolution is predominantly due to genetic drift. *Molecular Biology and Evolution* 27: 661-670.

Maynard Smith, J. 1966. Sympatric speciation. *American Naturalist* 100: 637-&.

Mayr, E. 1963. *Animal Species and Evolution*. Harvard University Press, Harvard.

Meyer, C. P. & Paulay, G. 2005. DNA barcoding: Error rates based on comprehensive sampling. *Plos Biology* 3: 2229-2238.

Mitra, S., Landel, H. & Pruett-Jones, S. 1996. Species richness covaries with mating system in birds. *Auk* 113: 544-551.

Moritz, C. & Cicero, C. 2004. DNA barcoding: Promise and pitfalls. *Plos Biology* 2: 1529-1531.

Morrow, E. H. & Pitcher, T. E. 2003. Sexual selection and the risk of extinction in birds. *Proceedings of the Royal Society B-Biological Sciences* 270: 1793-1799.

Morrow, E. H., Pitcher, T. E. & Arnqvist, G. 2003. No evidence that sexual selection is an 'engine of speciation' in birds. *Ecology Letters* 6: 228-234.

Møller, A. P. & Cuervo, J. J. 1998. Speciation and feather ornamentation in birds. *Evolution* 52: 859-869.

Panhuis, T. M., Butlin, R., Zuk, M. & Tregenza, T. 2001. Sexual selection and speciation. *Trends in Ecology & Evolution* 16: 364-371.

Peiponen, V. A. 1960. Verhaltensstudien am Blaukehlchen (*Luscinia s. svecica*). *Ornis Fennica* 37: 69-83.

Price, T. 2008. *Speciation in birds*. Roberts and Company Publishers, Colorado.

Questiau, S., Eybert, M. C., Gaginskaya, A. R., Gielly, L. & Taberlet, P. 1998. Recent divergence between two morphologically differentiated subspecies of bluethroat (Aves : Muscicapidae : *Luscinia svecica*) inferred from mitochondrial DNA sequence variation. *Molecular Ecology* 7: 239-245.

Questiau, S., Eybert, M. C. & Taberlet, P. 1999. Amplified fragment length polymorphism (AFLP) markers reveal extra-pair parentage in a bird species: the bluethroat (*Luscinia svecica*). *Molecular Ecology* 8: 1331-1339.

Qvarnström, A., Rice, A. M. & Ellegren, H. 2010. Speciation in *Ficedula* flycatchers. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365: 1841-1852.

Ratnasingham, S. & Hebert, P. D. N. 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes* 7: 10.

Rätti, O., Hovi, M., Lundberg, A., Tegelstrom, H. & Alatalo, R. V. 1995. Extra-pair paternity and male characteristics in the pied flycatcher. *Behavioral Ecology and Sociobiology* 37: 419-425.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 74: 5463-5467.

Schluter, D. 2009. Evidence for ecological speciation and its alternative. *Science* 323: 737-741.

Schluter, D. & Price, T. 1993. Honesty, perception and populaition divergence in sexually selected traits. *Proceedings of the Royal Society B-Biological Sciences* 253: 117-122.

Schulenburg, J., Hurst, G. D. D., Tetzlaff, D., Booth, G. E., Zakharov, I. A. & Majerus, M. E. N. 2002. History of infection with different male-killing bacteria in the two-spot ladybird beetle *Adalia bipunctata* revealed through mitochondrial DNA sequence analysis. *Genetics* 160: 1075-1086.

Seddon, N., Merrill, R. M. & Tobias, J. A. 2008. Sexually selected traits predict patterns of species richness in a diverse clade of suboscine birds. *American Naturalist* 171: 620-631.

Seehausen, O., Alphen, J. J. M. v. & Witte, F. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277: 1808-1811.

Seutin, G., White, B. N. & Boag, P. T. 1991. Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* 69: 82-90.

Sheldon, B. C. & Ellegren, H. 1999. Sexual selection resulting from extrapair paternity in collared flycatchers. *Animal Behaviour* 57: 285-298.

Slagsvold, T. & Lifjeld, J. T. 1986. Mate retention and male polyterritoriality in the pied flycatcher *Ficedula hypoleuca*. *Behavioral Ecology and Sociobiology* 19: 25-30.

Sokal, R. R. & Rohlf, F. J. 1995. *The Principles and Practice of Statistics in Biological Research,* 3 ed. W H Freeman & Co, New York.

Storchova, R., Reif, J. & Nachman, M. W. 2010. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. *Evolution* 64: 456-471.

Sæther, S. A., Sætre, G. P., Borge, T., Wiley, C., Svedin, N., Andersson, G. *et al.* 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* 318: 95-97.

Sætre, G. P., Borge, T., Lindell, J., Moum, T., Primmer, C. R., Sheldon, B. C. *et al.* 2001. Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Molecular Ecology* 10: 737-749.

Sætre, G. P., Borge, T., Lindroos, K., Haavie, J., Sheldon, B. C., Primmer, C. *et al.* 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. *Proceedings of the Royal Society of London Series B-Biological Sciences* 270: 53-59.

Sætre, G. P., Kral, M., Bures, S. & Ims, R. A. 1999. Dynamics of a clinal hybrid zone and a comparison with island hybrid zones of flycatchers (*Ficedula hypoleuca* and *F. albicollis*). *Journal of Zoology* 247: 53-64.

Sætre, G. P. & Sæther, S. A. 2010. Ecology and genetics of speciation in *Ficedula* flycatchers. *Molecular Ecology* 19: 1091-1106.

Taylor, E. B., Boughman, J. W., Groenenboom, M., Sniatynski, M., Schluter, D. & Gow, J. L. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* 15: 343-355.

Turner, G. F. 2002. Parallel speciation, despeciation and respeciation: implications for species definition. *Fish and Fisheries* 3: 225-229.

Veen, T., Borge, T., Griffith, S. C., Sætre, G. P., Bures, S., Gustafsson, L. *et al.* 2001. Hybridization and adaptive mate choice in flycatchers. *Nature* 411: 45-50.

Vortman, Y., Lotem, A., Dor, R., Lovette, I. J. & Safran, R. J. 2011. The sexual signals of the East-Mediterranean barn swallow: a different swallow tale. *Behavioral Ecology* 22: 1344-1352.

Webb, W. C., Marzluff, J. M. & Omland, K. E. 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Molecular Ecology* 20: 2390-2402.

Willows-Munro, S., Robinson, T. J. & Matthee, C. A. 2005. Utility of nuclear DNA intron markers at lower taxonomic levels: Phylogenetic resolution among nine *Tragelaphus* spp. *Molecular Phylogenetics and Evolution* 35: 624-636.

Wolfson, A. 1952. The cloacal protuberance: a means for determining breeding condition in live male passerines. *Bird Banding* 23: 159-165.

Yang, S., Ruuhola, T., Haviola, S. & Rantala, M. J. 2008. Effects of host-plant shift on immune and other key life-history traits of an eruptive Geometrid, *Epirrita autumnata* (Borkhausen). *Ecological Entomology* 33: 510-516.

Zink, R. M. & Barrowclough, G. F. 2008. Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* 17: 2107-2121.

Zink, R. M., Drovetski, S. V., Questiau, S., Fadeev, I. V., Nesterov, E. V., Westberg, M. C. *et al.* 2003. Recent evolutionary history of the bluethroat (*Luscinia svecica*) across Eurasia. *Molecular Ecology* 12: 3069-3075.

I

Ecology and Evolution

Open Access

Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*

Silje Hogner[1], Terje Laskemoen[1], Jan T. Lifjeld[1], Jiri Porkert[2], Oddmund Kleven[1,3], Tamer Albayrak[4], Bekir Kabasakal[4] & Arild Johnsen[1]

[1]National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172, Blindern, NO-0318, Oslo, Norway
[2]Gocarova 542, Hradec Kralove, 500 02, Czech Republic
[3]Norwegian Institute for Nature Research – NINA, P.O. Box 5685, Sluppen, 7485, Trondheim, Norway
[4]Department of Biology, Science and Art Faculty, University of Mehmet Akif Ersoy, Burdur, Turkey

Keywords
Assortative mating, mtDNA, reproductive isolation, sympatric divergence.

Correspondence
Silje Hogner, National Centre for Biosystematics, Natural History Museum, University of Oslo, Blindern 1172, NO-0318 Oslo, Norway. Tel: +47 22 85 18 86; Fax: +47 22 85 18 37; E-mail: silje.hogner@nhm.uio.no

Funding Information
Funding was provided by the Natural History Museum, University of Oslo (PhD fellowship to SH).

Received: 16 September 2012; Accepted: 17 September 2012

Ecology and Evolution **2012; 2(12): 2974–2988**

doi: 10.1002/ece3.398

Abstract

Mitochondrial DNA usually shows low sequence variation within and high sequence divergence among species, which makes it a useful marker for phylogenetic inference and DNA barcoding. A previous study on the common redstart (*Phoenicurus phoenicurus*) revealed two very different mtDNA haplogroups (5% K2P distance). This divergence is comparable to that among many sister species; however, both haplogroups coexist and interbreed in Europe today. Herein, we describe the phylogeographic pattern of these lineages and test hypotheses for how such high diversity in mtDNA has evolved. We found no evidence for mitochondrial pseudogenes confirming that both haplotypes are of mitochondrial origin. When testing for possible reproductive barriers, we found no evidence for lineage-specific assortative mating and no difference in sperm morphology, indicating that they are not examples of cryptic species, nor likely to reflect the early stages of speciation. A gene tree based on a short fragment of cytochrome c oxidase subunit 1 from the common redstart and 10 other *Phoenicurus* species, showed no introgression from any of the extant congenerics. However, introgression from an extinct congeneric cannot be excluded. Sequences from two nuclear introns did not show a similar differentiation into two distinct groups. Mismatch distributions indicated that the lineages have undergone similar demographic changes. Taken together, these results confirm that deeply divergent mitochondrial lineages can coexist in biological species. Sympatric mtDNA divergences are relatively rare in birds, but the fact that they occur argues against the use of threshold mtDNA divergences in species delineation.

Introduction

Many species exhibit high levels of intraspecific morphological and genetic variation. Variation in mitochondrial DNA (mtDNA) is particularly prevalent, due to its faster evolutionary rate compared with nuclear DNA (Avise 2000). Usually, such variation is confined to allopatric populations and can be explained by long periods of isolation with differing selection pressures and/or divergence due to genetic drift (Coyne and Orr 2004; Price 2008). High intraspecific mtDNA variation between individuals living in sympatry is less common, and more difficult to explain. Upon closer inspection, such divergent sympatric lineages often show evidence of divergence in other parts of the genome as well as reproductive isolation between the lineages, implying that they are in fact cryptic species (Hebert et al. 2004a; Haine et al. 2006). The concept of DNA barcoding, which applies the mitochondrial cytochrome c oxidase subunit 1 (COI) marker in animals, is based on the premise that there is low variation within species and large divergence gaps between sister species (Hebert et al. 2003, 2004b). Accordingly, provisional species are often suggested when sequence divergence exceeds a certain threshold (e.g., 10 times average intraspecific variation; Carr et al. 2011;

Kerr et al. 2007). However, there are examples of sympatric intraspecific divergences of a magnitude that exceeds normal sister species-level divergence (Wayne et al. 1990; Tominaga et al. 2009; Xiao et al. 2012), making species discrimination based on a strict divergence threshold in mtDNA too simplistic (Moritz and Cicero 2004).

In birds, deep sympatric mtDNA divergences have been found in a few species (e.g., Quinn 1992; Webb et al. 2011; Kerr et al. 2009b; Johnsen et al. 2010; Barrowclough et al. 2011). For example, common ravens (*Corvus corax*) show a 4% divergence between Holarctic and western North American lineages, with a high degree of sympatry and interbreeding (Webb et al. 2011) and males of the (*Manacus manacus*) collected from a single lek represented two groups with 3.5% divergence (Kerr et al. 2009b). The interpretation of such deep sympatric divergences is challenging and requires additional information about potential methodological pitfalls in mtDNA sequencing and reproductive barriers to gene flow in nuclear DNA.

Several hypotheses have been proposed to explain such high mtDNA variation (Webb et al. 2011). First, sympatric intraspecific divergences in mtDNA may be an artifact caused by nuclear mitochondrial pseudogenes, or numts. This is genetic material that has been translocated from the mitochondrial to the nuclear genome. These copies are assumed to be nonfunctional (Bensasson et al. 2001), evolve fast, accumulate frame shifts and stop codons, and show double peaks when sequenced (Bertheau et al. 2011; Triant and Hayes 2011). Second, as stated above, the divergent lineages may in fact reflect cryptic species, implying that the taxonomy is incorrect. Recent avian examples include thrushes (*Turdus spp.*) in western Amazonas (O'Neill et al. 2011), the winter wren (*Troglodytes troglodytes*) in North America (Toews and Irwin 2008), and seven nonpasserine migratory birds from the Philippines (Lohman et al. 2010).

A third possibility is that high mtDNA variation is caused by hybridization with a closely related species, which can lead to introgression of mtDNA (Coyne and Orr 2004; Bachtrog et al. 2006; Toews & Brelsford, 2012). Hybridization is common in birds, occurring in approximately one out of 10 species (Grant and Grant 1992; McCarthy 2006). However, as females are the heterogametic sex in birds, female hybrids are more likely to be affected by reduced viability and/or fertility than males (Haldane 1922), which reduces the likelihood of introgression of the maternally inherited mtDNA. Nevertheless, Taylor et al. (2011) found evidence for hybridization between the sister species common murre (*Uria aalge*) and thick-billed murre (*U. lomvia*), with mtDNA introgression from the thick-billed murre into the common murre. Another example is mtDNA introgression between the golden-winged warblers (*Vermivora chrysoptera*) and the blue-winged warblers (*V. cyanoptera*) in North America (Shapiro et al. 2004).

Fourth, deep mtDNA divergence can reflect long periods of geographical isolation followed by secondary contact. The divergence might be a result of neutral differences within a single species, and thus represent a historical artifact of divergent lineages that have remerged (Webb et al. 2011). In the absence of reproductive barriers, such remerging lineages will be expected to collapse into one (speciation in reverse). The fixation of ancestral allelic lineages can be due to either drift or selection, and it produces a reciprocally monophyletic gene tree (Neigel and Avise 1986). If the two populations have been separated long enough, with little or no gene flow between them, they may have accumulated genetic and phenotypic differences, which might result in reproductive barriers in the form of different morphological, physiological or behavioral traits (Coyne and Orr 2004). Reproductively isolated forms might thus arise if local adaptations are strong, colonization of alternative habitats is eliminated and reproductive contact is reduced (Nosil et al. 2005; Sobel et al. 2010). If secondary contact later occurs, a shift in mate recognition systems and mate preferences may lead to assortative mating (precopulatory barrier) or gamete incompatibilities (postcopulatory, prezygotic barrier) as a result of the earlier allopatry, and the genetic variation between the two populations will be maintained (early speciation). Sperm morphology has a genetic basis (Birkhead et al. 2005), shows remarkable levels of diversification (reviewed in Pitnick et al. 2009) and has shown geographical variation in some avian studies (Lüpold et al. 2011; Schmoll and Kleven 2011). Differences in sperm morphology may thus contribute to prezygotic reproductive barriers in the early stages of speciation (Coyne and Orr 2004). Finally, deep mtDNA divergence may reflect maintenance of two or more ancestral lineages in a panmictic population with large effective population size (Avise et al. 1988; Webb et al. 2011).

In a recent DNA barcoding study, Johnsen et al. (2010) found two different COI lineages in the common redstart. The divergence between these two haplotype lineages (hereafter referred to as haplogroups) was in the magnitude of 5%, suggesting that these lineages separated about 2 million years ago according to the conventional molecular clock estimate (Bromham and Penny 2003; but see Pulquério and Nichols 2007; Weir and Schluter 2008). These two haplogroups were initially found to interbreed in one mixed pair from Norway (Johnsen et al. 2010). Our main aims in the present study are twofold. First, we describe the distribution of the two haplogroups found in Johnsen et al. (2010) across the breeding range of the common redstart, and hence examine their degree of sympatry

and interbreeding in detail. Second, we explore the above hypotheses for how this deep mtDNA variation may have originated, combining sequence data from two mtDNA regions (control region and COI) and two nuclear Z-linked introns (BRM-15 and ALDOB-6), with data on degree of assortative mating and sperm size variation between the haplogroups. We test the following predictions related to each hypothesis. (1) From the numt hypothesis, we predict to find stop codons and double peaks in the sequences. (2) From the cryptic species hypothesis, we predict to find reproductive barriers, such as assortative mating or differences in the sperm morphology, and divergence in nuclear DNA that is related to the divergence in mtDNA. (3) From the recent hybridization hypothesis, we predict that one of the haplogroups would cluster together with one of the other extant *Phoenicurus* species. (4) From the geographic isolation hypothesis, we predict that there will be structure in the geographical distribution of the two haplogroups, and that they will show different mismatch distributions due to different demographic histories. If the lineages are in the process of remerging (speciation in reverse), there should be no reproductive barriers and little or no structure in the nuclear sequences with respect to the mtDNA lineages yet high nuclear nucleotide variation, whereas if they are in the process of further divergence (early speciation), we would predict to find some evidence for reproductive isolation and a pattern of divergence in the nuclear data as a result of the two lineages being effectively separated. Finally, (5) from the coexistence in one panmictic population hypothesis, we predict lack of geographic structure and reproductive barriers, similar mismatch distributions, and no divergence in the nuclear introns.

Materials and Methods

Study species

The common redstart is a small (~15 g), sexually dimorphic passerine bird, breeding in the Western Palearctic (Fig. 1), and wintering in North Africa. The breeding system is predominantly social and genetic monogamy (Kleven et al. 2007), but instances of polygyny have been observed (del Hoyo et al. 2005).

Samples

High quality DNA

Breeding redstarts from Norway, the Czech Republic, Finland, Morocco, Mongolia, Iran, Spain, and Turkey were caught at their respective breeding grounds during spring 2002, 2006, 2009, 2010, and 2012. We collected up to 25-μl blood by brachial venipuncture and stored the blood in 96% ethanol. In addition, blood samples were collected from migratory birds from Israel (see Table 1 for sample details). Birds were caught using mist nets and playback in the beginning of their breeding season in their breeding

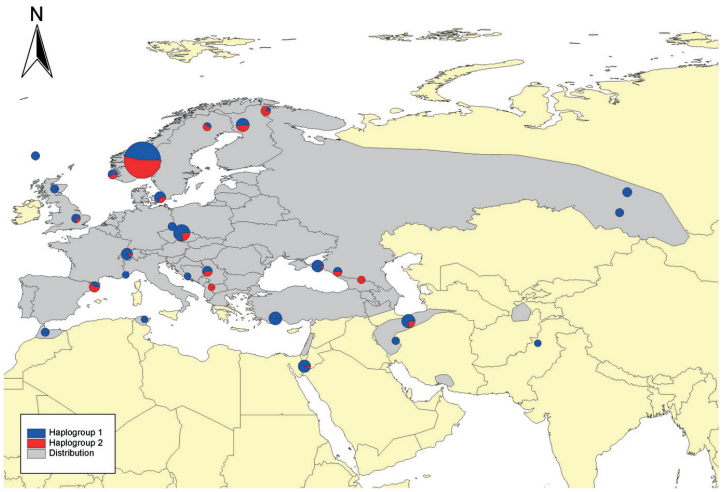


Figure 1. Map showing the distribution of the common redstart (shaded area), and the sampling locations (circles) with the relative frequency of the two haplogroups (blue = haplogroup 1 and red = haplogroup 2). Breeding birds are collected from all locations, except from Israel (migrating birds). On the basis of all common redstart samples (N = 387), both sequenced for the long COI and the short COI fragment.

Table 1. Basic sample information for (a) contemporary common redstarts, where DNA was extracted from either fresh blood, tissue or feather samples and subjected to long COI sequencing, and (b) up to 150-years old museum skins (and one contemporary population from Finland) where DNA was extracted from foot pads or dried blood samples (Finland) and subjected to short COI sequencing.

Locality	Subspecies	Latitude	Longitude	*n* m/f/u*	*N* Sperm	Haplotype 1/2	Year
a)							
Norway, south-east	*phoenicurus*	61°16′N	12°17′E	100/78/19	67	94/103	2002–2010
Norway, west	*phoenicurus*	59°18′N	04°52′E	2/2/0	0	2/2	2005
Czech Republic	*phoenicurus*	50°11′N	15°55′E	13/9/9	8	24/7	1991–2010
Russia, Adygea	*phoenicurus*	44°11′N	40°04′E	3/1/0	0	2/2	2004
Russia, Krasnodarskiy Kray	*phoenicurus*	45°04′N	36°59′E	6/5/0	0	11/1	2004–2006
Serbia	*phoenicurus*	44°16′N	19°53′E	0/0/7	0	4/3	2008
Spain	*phoenicurus*	41°50′N	02°23′E	7/1/0	0	3/5	2012
Turkey	*samamisicus*	36°58′N	30°26′E	1/0/15	0	16/0	2010
Iran	*samamisicus*	36°32′N	51°3′E	14/4/0	0	13/5	2010
Morocco	*phoenicurus*	33°32′N	5° 6′W	2/0/0	0	2/0	2008
Israel	*samamisicus*	29º33′N	34º56′E	0/0/14	0	13/1	2007–2008
b)							
Norway, north	*phoenicurus*	69°4′N	28°55′E	0/6/0	0	1/5	1866–1966
Finland	*phoenicurus*	66°54′N	25°22′E	0/17/0	0	9/8	1992–2006
Sweden	*phoenicurus*	66°38′N	19°51′E	3/0/0	0	1/2	1915–1970
Faroe Islands	*phoenicurus*	62°00′N	40°24′W	3/0/0	0	3/0	1898–1910
Norway, south-east	*phoenicurus*	61°16′N	12°17′E	2/0/3	0	4/1	1882–1908
Norway, west	*phoenicurus*	58°41′N	05°34′E	0/1/0	0	1/0	1908
Scotland	*phoenicurus*	57°03′N	3°03′W	3/0/0	0	3/0	1915–1919
Denmark	*phoenicurus*	55°42′N	12°34′E	10/0/0	0	7/3	1890–1977
England	*phoenicurus*	51°46′N	40°10′W	4/0/0	0	3/1	1933–1956
Germany	*phoenicurus*	51°10′N	14°26′E	3/0/0	0	3/0	1893–1952
Switzerland	*phoenicurus*	46°56′N	7°26′E	10/0/0	0	9/1	1930–1953
France	*phoenicurus*	43°42′N	7°14′E	1/0/0	0	1/0	1937
Karaku, Pakistan	*samamisicus*	33°6′N	71°5′E	0/1/0	0	1/0	1876
Croatia	*phoenicurus*	43°30′N	16°55′E	1/0/0	0	1/0	unknown
Russia, Barnaul	*phoenicurus*	53°19′N	83°46′E	3/0/0	0	3/0	1896–1933
Russia, Caucasus	*phoenicurus*	42°55′N	43°45′E	2/0/0	0	0/2	unknown
Russia, Tomsk	*phoenicurus*	56°28′N	84°57′E	5/0/0	0	5/0	1896–1923
Macedonia	*phoenicurus*	41°47′N	20°32′E	1/0/0	0	0/1	1935
Turkey	*samamisicus*	36°58′N	30°26′E	1/0/0	0	1/0	1876
Tunisia	*phoenicurus*	36°49′N	10°09′E	1/0/0	0	1/0	1938
Iran	*samamisicus*	36°32′N	51°3′E	0/0/5	0	3/2	2010
Morocco	*phoenicurus*	33°32′N	5°6′W	1/0/0	0	1/0	1919

*m, males; f, females; u, unknown.

territories, or with a clap-net (with a meal-worm for a bait) during the period of feeding the chicks. All birds were putatively unrelated, predominantly adult breeders ($N = 254$), but in some cases ($N = 22$) where the female could not be sampled, we instead sampled one chick per nest as a representative of the female mtDNA lineage. In addition, tissue samples collected from breeding birds in Russia (2004–2006, $N = 16$) and Serbia (2008, $N = 7$) were provided by Yale Peabody Museum (New Heaven) and tissue samples of one Moussier's redstart (*P. moussieri*), one blue-fronted redstart (*P. frontalis*), and two Eversmann's redstart (*P. erythronotus*), were provided from the Natural History Museum of Copenhagen (see SI Table 1 for more information).

Degraded DNA

Feather samples were collected from nine females in Finland in June 1992–1994 and blood samples were collected (and dried) from nine unrelated chicks during spring 1998 and 1999 in Hradec Kràlové, Czech Republic. Toe-pad samples were collected from 88 adult museum specimens (*P. phoenicurus*) from the Natural History Museum of London (Tring), Harrison Museum (London), the Natural History Museum of Oslo, and the Natural History Museum of Copenhagen. Of these, the majorities were adult breeders, except three individuals collected in Copenhagen, Denmark and three individuals sampled in the Faroe Islands during migration (SI table 2). Finally,

toe-pad samples were collected from Prezevalski's redstart (*P. alaschanicus*, $N = 2$), blue-capped redstart (*P. caeruleocephalus*, $N = 2$), blue-fronted redstart ($N = 2$), Hodgson's redstart (*P. hodgsoni*, $N = 2$), Moussier's redstart ($N = 2$), and white-tailed redstart (*P. schisticeps*, $N = 2$) in order to investigate the relationship between the two *P. phoenicurus* haplogroups and closely related species of the genus *Phoenicurus* (samples were collected from the Natural History Museum of London (Tring), Yale Peabody Museum (New Heaven), Natural History Museum of Copenhagen and Natural History Museum of Oslo. SI Table 2).

Genbank sequences

COI sequences from six black redstart (*P. ochruros*), five Eversmann's redstart, three Güldenstädt's redstart (*P. erythrogastrus*), and five daurian redstart (*P. auroreus*) were downloaded from Genbank (see SI Table 3 for more information).

DNA extraction and PCR

DNA from the blood samples was extracted following the protocol for the E.Z.N.A blood kit (Omega Bio-Tek, Inc, Norcross, Georgia). DNA from the toe-pads was extracted following the protocol form E.Z.N.A-tissue Kit (Omega Bio-Tek, Inc, Norcross, Georgia), or using a Mole-extraction robot following the manufacturers' protocol (Mole Genetics AS, Norway).

For the common redstart, two fragments of mtDNA (COI, 700bp and control region, 421bp), and two Z-linked introns (BRM-15, 311bp, and ALDOB-6, 531bp) were sequenced (Genbank accession number JX945383-JX945521). COI was sequenced for a total of 201 redstarts from different populations (Table 1: details and sequences also available at the BOLD website (http://www.barcodinglife.com/), project NorBOL – Birds – Phoenicurus). For the samples consisting of degraded DNA and some of the high quality DNA samples, a short piece of the COI (120bp) was amplified and a restriction enzyme (Aci II) was used to determine the haplogroup ($N = 88$ for degraded DNA and $N = 98$ for high quality DNA). The restriction enzyme was chosen so that it would cut in one of the conserved sites in haplogroup 2, making two bands visible on an electrophoresis gel for this haplogroup. For the other haplogroup, only one band was visible. Ten individuals were also analyzed using both restriction cutting and sequencing in order to confirm the validity of the methods. In order to get the exact haplotype for the Z-linked introns, a total of 54 females of the common redstart were chosen from the south-east population in Norway (of these, only 42 worked for both introns chosen herein). The same 54 individuals were also sequenced for the control region. For the other *Phoenicurus* samples, the long COI ($N = 4$) or short COI ($N = 12$) fragment was sequenced (project NorBOL – Birds – Phoenicurus). All regions were amplified in PCR reaction volumes of 10-μL, containing dH_2O, 1X PCR buffer II (Applied Biosystems, Foster City, California), 1.5-mM magnesium, 0.2-mM dNTP (ABgene, Epsom, UK), 0.5-mM forward and reverse primer, 3% Dimethyl sulfoxide (DMSO), 0.25 U AmpliTaq DNA polymerase (Applied Biosystems), and approximately 50-ng DNA template. The amplifications were run on a DNA Engine Tetrad 2 (MJ Research, Waterton, MA, USA). The following profile was used: 95° C for 1 min, 94°C for 30 sec, primer-specific annealing temperature (see Table 2) (55–60°C) for 30 sec, 72°C for 1 min, then the second to forth step another 34–39 cycles before the last step, 72°C for 10 min. A 3-μl PCR-product was electrophoresed in 1% agarose TBE to confirm amplification success and to exclude any contamination.

The remaining PCR-product was purified by digesting unincorporated nucleotides and primers using diluted (1:9) ExoSap-It (United States Biochemical, Cleveland, Ohio) run at 37°C for 45 min followed by 80°C for

Table 2. Primer information and amplification conditions.

Locus	Class[1]	Primer sequence (5′-3′)	PCR[2]	Reference
Aldob-6	Z	F: AGACCATGATCTCCAGCGCT R: CCTTCCAGGTAGACATGATG	56	Borge et al. 2005
Brm-15	Z	F: AGCACCTTTGAACAGTGGTT R: TACTTTATGGAGACGACGGA	56	Borge et al. 2005
COI-ExtF BirdR2	m	F: ACGCTTTAACACTCAGCCATCTTACC R: ACTACATGTGAGATGATTCCGAATCCAG	55	Johnsen et al. 2010
PhSa-F1 PhSa-R1	m	F: AACGTAGTCGTCACAGCCCATGCTT R: TTATTCGRGGRAATGCTATG	55	This study
L437 H1248	m	F: CTCACGAGAACCGAGCTACT R: CATCTTCAGTGTCATGCT	52	Tarr 1995

[1]DNA class: Z, Z-linked; m, mtDNA.
[2]Annealing temperature.

15 min to inactivate the enzyme. The PCR products were then sequenced using BigDye Terminator sequencing buffer and v 3.1 Cycle Sequencing kit (Applied Biosystems). The sequences were aligned and edited using ClustalW in the program Mega v 5 (Tamura et al. 2007) and modified manually. Each base was called, using both forward and reverse sequencing reads for each strand. All sequences for each locus were adjusted to the same length as the shortest sequence of that locus for comparison.

Genetic analyses

Molecular gene trees were constructed using the neighbor-joining method implemented in Mega v 5 (Tamura et al. 2007), using the Kimura two-parameter model and 10,000 bootstrap replicates.

In order to examine the genetic structure of the redstart populations, analysis of molecular variance (AMOVA) was run using the program Arlequin v 3.5 (Excoffier et al. 2005). Pairwise population differences were estimated using the F_{ST} statistic (Weir and Cockerham 1984) implemented in Arlequin, with default settings for the population comparisons. In these analyses, we included seven populations with minimum seven individuals sequenced for the long COI fragment. Sequential Bonferroni correction was applied to adjust critical *P*-values for multiple statistical testing (Rice 1989). DNAsp v 5 (Librado and Rozas 2009) was used to calculate nucleotide variation, π, of the Z-introns (Hudson et al. 1987).

To test for historical demographic events within the two haplogroups (Johnsen et al. 2010), we first calculated Tajima's *D* (Tajima 1989) for the concatenated COI and control region sequences using DNAsp v 5. The sign of the test statistic can indicate a recent bottleneck (positive Tajima's *D*) or population expansion (negative Tajima's *D*). Second, we compared the observed frequency distribution of pairwise nucleotide differences among individuals within each of the haplogroups with the expected distribution from a sudden population expansion (mismatch distribution), using Arlequin v 3.5. If a population has experienced a long lasting demographic equilibrium or a decline, then a multimodal distribution should be displayed, whereas a unimodal distribution should be displayed if a population has experienced a sudden demographic expansion (Slatkin and Hudson 1991; Rogers and Harpending 1992). However, recent changes are not always detectable in a mismatch distribution, because they might be masked by threshold effects, time lags or earlier demographic events (Rogers and Harpending 1992; Lavery et al. 1996). Arlequin tests the goodness-of-fit to this model using SSD test statistics (the sum of squared differences between the observed and the estimated mismatch distributions; Rogers and Harpending 1992).

We used the general nonlinear least-square approach to estimate the demographic mutation time parameter tau, $\tau = 2\mu t$, where μ is the mutation rate per generation of the DNA fragment and t is the number of generations. Assuming a generation time of 1 year and the standard molecular clock of mtDNA divergence of 2% per million years (Bromham and Penny 2003; but see Weir and Schluter 2008; Lande et al. 2003), we estimated the time since expansion of the two haplogroups using the formula above (see Sætre et al. 2012 for further details).

Assortative mating

Generalized linear models (GLZ) with binomial distribution were performed in Statistica, to test for assortative mating between the two haplogroups among breeding pairs. We had data from 68 pairs, from Norway ($N = 60$) and Czech Republic ($N = 8$), respectively. However, as there were no haplogroup two females represented in the Czech population, testing for assortative mating would be noninformative and thus this population was excluded.

Sperm measurements

We obtained sperm samples from 67 males from three subpopulations in south-east Norway (Røros N 62º 37′, E 11º 38′, Trysil N 61º16′, E 12º 17′ and Aurdal N 60º 39′, E 9º37′) and 8 males from the Czech Republic (Hradec Králové N 50º 11′, E 15º 55′). Sperm samples were collected by gently massaging the cloacal protuberance of breeding males using a similar technique as described in Wolfson (1952). The ejaculate was collected using a microcapillary tube and fixed in a 5% formalin solution. Sperm morphology data were obtained for each individual, from 10 normal and undamaged sperm, as 10 sperm provides an accurate estimate of each individual's sperm length (Laskemoen et al. 2007). The following measurements were obtained (±0.1 μm); head length, midpiece length, tail length, flagellum length and total length, where flagellum length is the sum of midpiece + tail length, and total length the sum of head + midpiece + tail. For each sperm trait, we used the means within individuals. All measurements were obtained from digital images captured at a magnification of 160× using a Leica DFC420 camera mounted on a Leica DM6000 B digital light microscope (Leica Microsystems, Switzerland). To avoid observer effects, one person (T.L.) conducted all sperm measurements. All sperm components were normally distributed (Shapiro–Wilks tests, all $W > 0.96$, all $P > 0.05$). Statistical analyses of assortative mating and differentiation in sperm morphology were conducted using Statistica v 7.1 (StatSoft Inc).

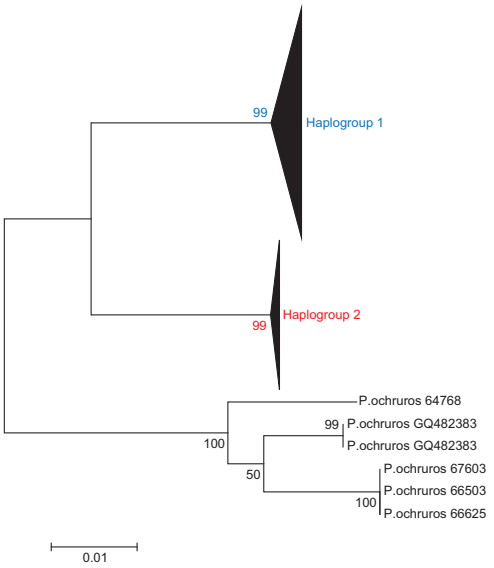


Figure 2. Neighbor-joining tree (K2P, 10,000 bootstrap replicates) based on COI (545bp) for contemporary common redstarts (N = 201), with black redstart as outgroup. The two common redstart haplogroups consist of 122 and 79 individuals, respectively. Only bootstrap values above 50% are shown. Blue = haplogroup 1 and red = haplogroup 2.

Results

Haplogroup characterization and distribution

A gene tree based on the COI region from different populations (Norway, Czech Republic, Iran, Turkey, Morocco, Russia, Spain, and Serbia) with black redstart as an outgroup showed that the COI region consists of two clearly separated groups (Fig. 2), supported by high bootstrap values (98% and 99%). There is some variation within each of these two groups (SI Fig. 1). The split (~5%) between these two haplogroups suggests that the haplogroups separated about 2 million years ago, assuming a standard avian molecular clock of 2% sequence divergence per million years (e.g., Päckert et al. 2007). We also found two haplogroups for the control region (using the south-east Norwegian population, data not shown), which matched respective groups in the COI as would be expected for two regions in the mt genome. A neighbor-joining tree based on concatenated sequences of the two mtDNA regions is shown in Figure 3.

The two clades coexist in Scandinavia, Great Britain, and central to eastern Europe (Fig. 1). There is an overrep-

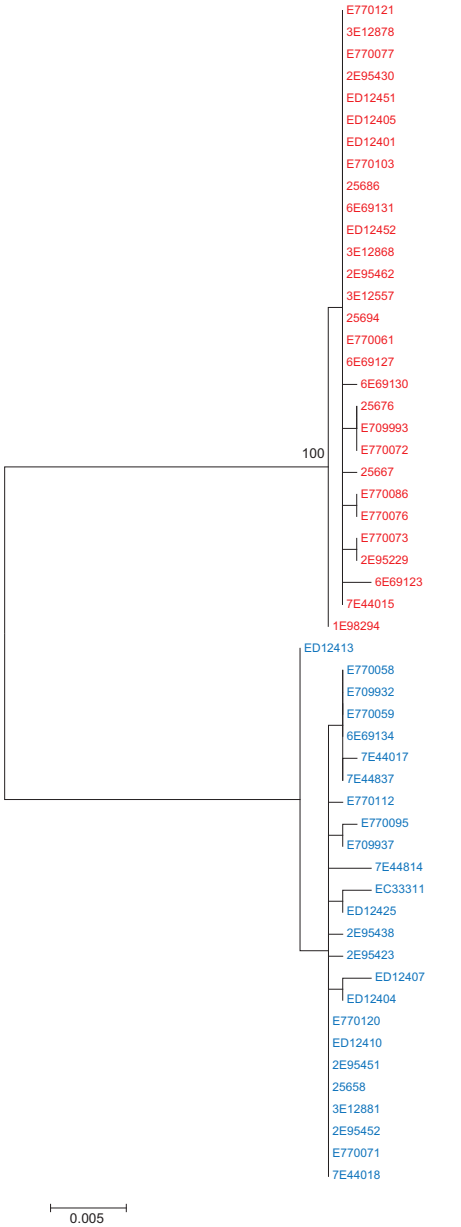


Figure 3. Neighbor-joining tree (K2P, 10,000 bootstrap replicates) based on the COI and control region (1121bp) combined for 54 common redstart females from the south-east Norwegian population. Only bootstrap values above 50% are shown. Blue = haplogroup 1 and red = haplogroup 2.

Table 3. Pairwise F_{st} (below diagonal) with *P*-values (above diagonal), for contemporary breeding populations. Bold = significant after sequential Bonferroni correction.

	Norway	Czech Republic	Serbia	Russia	Iran	Turkey	Spain
Norway		**<0.001**	0.77	0.03	0.02	**<0.001**	0.48
Czech Republic	0.22		0.12	0.38	0.09	**<0.001**	0.01
Serbia	−0.09	0.13		0.58	0.34	0.01	0.36
Russia	0.11	0.00	−0.06		0.43	0.01	0.03
Iran	0.10	0.10	−0.07	−0.03		0.02	0.03
Turkey	0.33	0.17	0.45	0.14	0.16		**<0.001**
Spain	−0.03	0.41	−0.07	0.21	0.17	0.63	

resentation of haplogroup 1 in western Europe and southern parts of the breeding range, whereas haplogroup 2 is mainly located in north and eastern Europe, and usually in coexistence with haplogroup 1. The AMOVA revealed significant differentiation in COI among seven contemporary breeding populations (F_{ST} = 0.17, $P < 0.001$). Pairwise comparisons showed significant differentiation between four of the populations (Norway vs. Czech R., and Turkey vs. Norway, Czech R. and Spain; Table 3).

When analyzing Z-intron variation among 42 females from the Norwegian population, we found no distinct groups related to the ones found in the mtDNA analyses (Fig. 4a,b). This is expected, as the Z-introns and mtDNA would segregate independently within a single population. We found the combined nucleotide diversity, the π-value, for the two introns to be 0.00614 (± 0.00043), which is the second highest π-value among 22 passerines sequenced in our lab (0.00079–0.00621).

A NJ tree based on the short COI sequence (120bp) from all 11 *Phoenicurus* species shows that the two haplogroups within the common redstart are unique (Fig. 5). Furthermore, in a dataset of long COI sequences (544bp) from seven species, including all the ones that appear to be most similar to the common redstart in the short COI tree (Fig. 5), the two haplogroups cluster together and are clearly different from the other extant *phoenicurus* species (SI Fig. 2).

Demographic patterns

Tajima's *D* estimates for the COI and control region combined were significantly negative for both haplogroups (haplogroup 2: Tajima's *D*: −1.855, $P < 0.05$; haplogroup 1: Tajima's *D*: −1.863, $P < 0.05$), and consistent with a population expansion (ArisBrosou and Excoffier 1996). A sudden population expansion was further supported using the mismatch distribution analysis, as both haplogroups fitted this model (Fig. 6a,b). Estimates of the time since the sudden expansion for the two haplogroups, suggest that they both expanded relatively recently, haplogroup 1: t = 46,025 years ago (90% CI: 29,040, 65,430), haplogroup 2: t = 26 020 years ago (90% CI: 11,357, 40,255).

Tests of reproductive barriers

In the Norwegian population, 35 of 60 pairs consisted of individuals belonging to the same haplogroup (Fig. 7), whereas in the Czech population, four of six sampled pairs consisted of same haplogroup individuals. We found no significant departure from random mating (N = 60 pairs, Wald = 1.54, P = 0.21) with respect to haplotype. Furthermore, there were no differences in any of the sperm traits between the two haplogroups (Table 4), and no significant differences in any of the sperm traits among the different populations (all $F_{3,71} < 0.67$, all $P < 0.57$).

Discussion

Our results confirm those found in Johnsen et al. (2010), that there are two highly divergent, coexisting mtDNA haplogroups in the common redstart. The two haplogroups show some geographic structure, with haplogroup 1 occurring all over the species distribution, whereas haplogroup 2 occurs predominantly in Northern Europe and parts of Western Asia, an area in which the two lineages are sympatric and interbreed to a large extent. Variation at two Z-linked introns was not related to mtDNA variation. When testing for possible reproductive barriers, we found no evidence for assortative mating and no differentiation in sperm morphology between the two haplogroups.

The magnitude of the divergence within the common redstart (5%) exceeds the divergence found in mtDNA between many sister species (Tavares and Baker 2008). Such deep, sympatric splits have only been found in a handful of other bird species (e.g., Quinn 1992; Webb et al. 2011). There are several possible explanations for the origin and maintenance of such high mitochondrial diversity. First, seemingly high variation in the mtDNA has been shown to sometimes be a result of nuclear

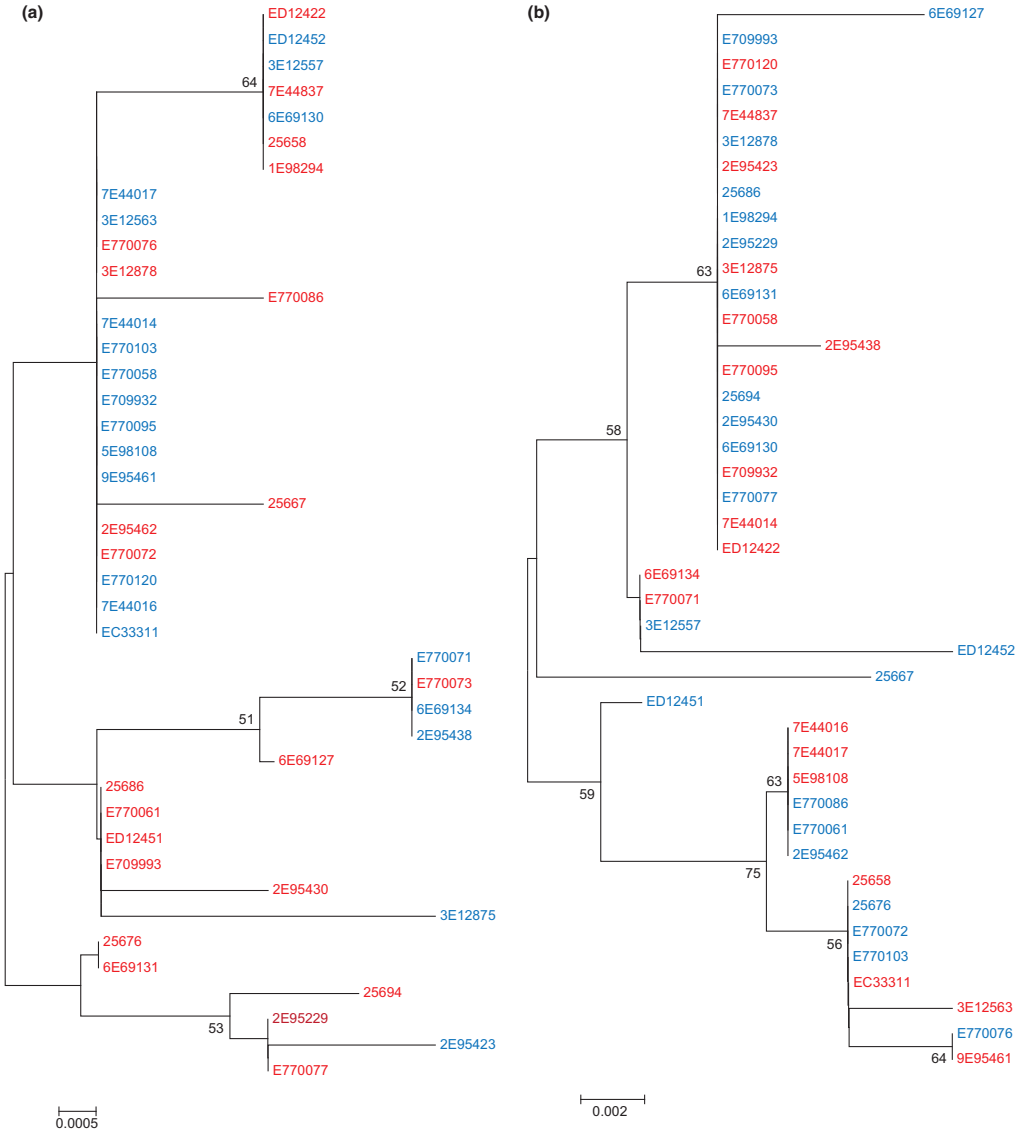


Figure 4. Neighbor-joining tree (K2P, 10,000 bootstrap replicates) based on two nuclear introns (a) = ALDOB-6 and (b) = BRM-15 for 42 common redstart females from the Norwegian south east population. Only bootstrap values above 50% are shown. Blue = haplogroup 1 and red = haplogroup 2.

mitochondrial pseudogenes, numts (Bensasson et al. 2001). To test for numts, we searched for stop codons and double peaks in the COI region, and double peaks in the control region sequences, and found no evidence for this. These two mtDNA fragments, in addition to 16S sequenced by Johnsen et al. (2010), cover a substantial proportion of the mtDNA, suggesting that the two haplogroups found here are not numts. In addition, Johnsen

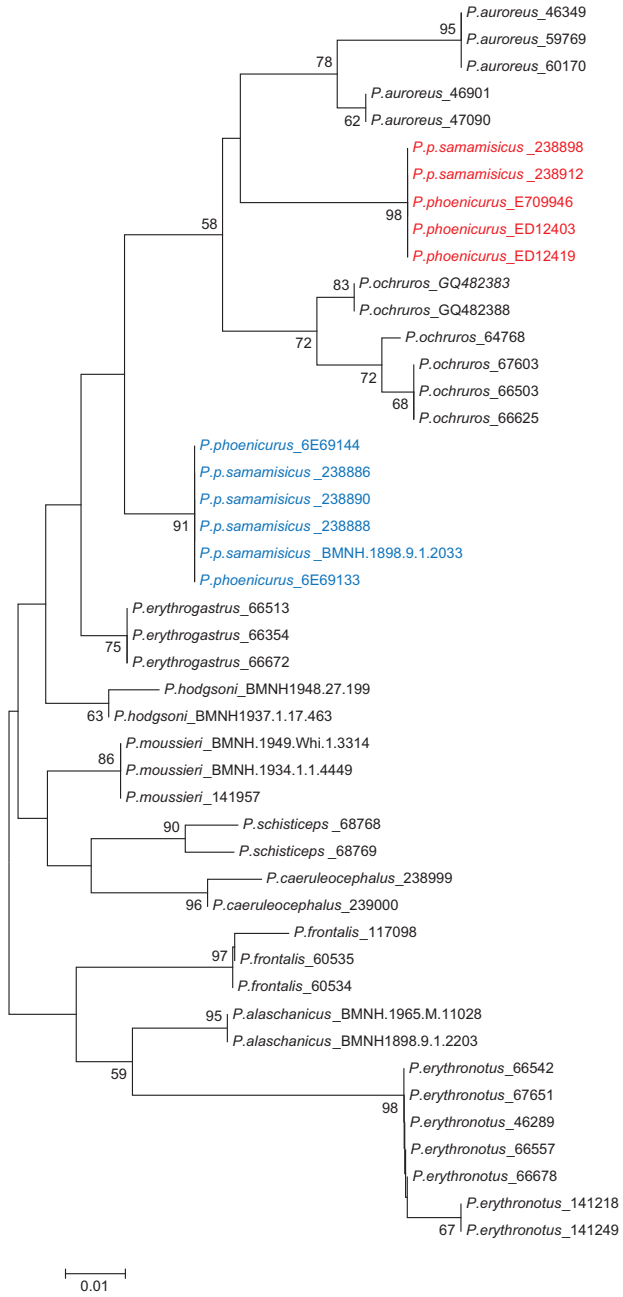


Figure 5. Neighbor-joining tree (K2P, 10,000 bootstrap replicates), based on a short fragment of COI (120bp) for all *Phoenicurus* species. Only bootstrap values above 50% are shown. Blue = haplogroup 1 and red = haplogroup 2.

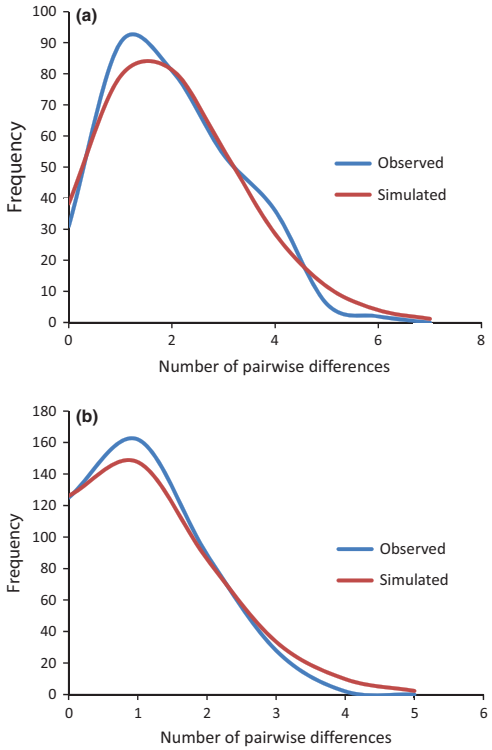


Figure 6. Mismatch distributions based on the combined alignment of the COI and control region for the 54 redstart females from the Norwegian south east population, for (a) haplogroup 1, and (b) haplogroup 2.

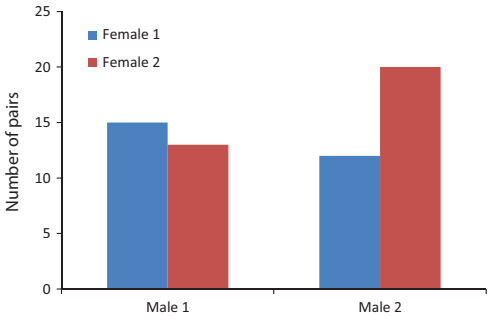


Figure 7. The proportion of common redstart individuals mating with their own, and the opposite haplogroup ($N = 60$ pairs). Blue = haplogroup 1 and red = haplogroup 2.

et al. (2010) ran a XL-PCR, which also supported mitochondrial origin of both haplogroups.

Second, with such a high level of divergence, the two haplogroups might represent cryptic species. However, there was no indication that this is the case, as we found no evidence for assortative mating with respect to haplogroups, no consistent divergence in nuclear introns and no differentiation in sperm morphology. Our results are similar to a study of the common raven by Webb et al. (2011), in which they found a high degree of mixing between two distinct mtDNA clades and no relationship with phenotype. Other studies have found indications of cryptic species based on deep splits in mtDNA, e.g., winter wrens in North America, where song differs and assortative mating was evident (Toews and Irwin 2008). Proponents of DNA barcoding have advocated the use of a threshold level of mtDNA divergence to delimit species (Hebert et al. 2004a,b). Indeed, several recent bird studies using DNA barcoding have suggested provisional species based on such a threshold (<2.5% Kerr et al. 2007, 2009a). Our study shows that using a threshold level to define species may sometimes lead to wrong conclusions, as the common redstart has a divergence of 5%, but is clearly just one biological species. This supports previous critiques of the threshold species concept (Moritz and Cicero 2004). However, as the two haplogroups form a monophyletic group (see SI Fig. 2) that is distinct from other closely related *phoenicurus* species, barcoding can still be used for species identification for this species.

Third, high mtDNA divergence, and inconsistency between mtDNA and nuclear gene trees, can also be a result of hybridization between closely related species, with introgression of mtDNA from one species to the other (Shapiro et al. 2004; Taylor et al. 2011). Common redstarts are known to hybridize with the black redstart in Central Europe, even giving rise to apparently fertile hybrids (Grosch 2004). However, given that none of the two haplogroups matched any of the mtDNA haplotypes found for other extant congenerics, the high divergence in common redstarts seems unlikely to be a result of introgression from other extant *Phoenicurus* species (see also SI Fig. 2). However, we cannot exclude the possibility of introgression of mtDNA from an extinct congeneric or an unsampled extant lineage.

The fourth, and perhaps most parsimonious, explanation of the occurrence of two distinct haplogroups, is that they arose in geographically isolated refugia during previous glaciation periods in Eurasia, and later came into secondary contact. Geographical isolation with secondary contact would predict geographical structure among the mtDNA haplogroups with concomitant differences in demographic history and high nucleotide variation in nuclear introns. The Z-intron nucleotide variation found in this study is the second highest among 22 recently analyzed passerine

Table 4. Sperm morphology of males from the two common redstart haplogroups, with the corresponding ANOVA statistics.

	Haplogroup 1 (N = 43)	Haplogroup 2 (N = 32)	ANOVA	
Sperm trait	Mean ± SD	Mean ± SD	$F_{1,73}$	P
Head	17.8 ± 0.9	17.8 ± 0.8	0.04	0.85
Midpiece	129.0 ± 6.0	127.5 ± 5.6	1.58	0.21
Tail	17.4 ± 2.6	17.4 ± 2.7	0.003	0.96
Total	164.4 ± 5.0	162.6 ± 6.6	1.75	0.19
Flagellum	146.6 ± 5.0	144.9 ± 6.5	1.69	0.20

species (Hogner et al. 2012; Gohli et al., unpubl. data; this study). We found significant genetic differentiation in a subset of seven populations, revealing some geographic structuring, with haplogroup 1 found throughout the whole geographic distribution area, and haplogroup 2 found predominantly in Northwestern Europe and parts of Western Asia (see Fig. 1). On the other hand, the two haplogroups showed similar demographic patterns, with evidence for sudden population expansion during the recent evolutionary history (<50,000 years ago). However, using the conventional molecular clock estimate of 2% divergence per million years (Bromham and Penny 2003; Päckert et al. 2007; but see Pulquério and Nichols 2007; Weir and Schluter 2008), the magnitude of the genetic distance of the haplogroups suggest that they originated in late Pliocene/early Pleistocene (more than 2 million years ago). Thus, the recent glaciation periods in Eurasia might have played a minor role for the origin of the two mtDNA haplogroups observed in the common redstart today. Furthermore, it is difficult to identify the possible location of such old refugia from the present data. The common redstart is taxonomically divided into two subspecies, *P. phoenicurus* and *P. samamisicus*, with *samamisicus* being confined to south-eastern Europe and south-western Asia, including Caucasus, which is a well-known glacial refugium (Hewitt 2000). Hence, the subspecies differentiation possibly reflects a period of geographic isolation, with those evolving into *samamisicus* residing in Caucasus. However, this separation into two subspecies is not related to the divergent mtDNA lineages, as we found a mixture of the two subspecies in the two mtDNA haplogroups (see Fig. 5). This suggests that the isolation event leading to the two divergent mtDNA lineages occurred considerably earlier than the isolation that gave rise to morphological differences between the two subspecies. Alternatively, the mtDNA haplogroups and the subspecies result from the same vicariance event where interbreeding has allowed neutral mtDNA introgression at the same time as selection has maintained the morphological differences.

When two lineages become separated, several factors, including taxon-specific rates of genetic differentiation, the severity of range reduction and timing of allopatric isolation, will play important roles in determining whether these lineages will become reproductively isolated from each other or not (Zamudio and Savage 2003). We found no evidence for reproductive isolation between the haplogroups, neither in the form of assortative mating (a possible precopulatory barrier) or in sperm morphology (a possible postcopulatory, prezygotic barrier; Coyne and Orr 2004). This is similar to the pattern found in the common raven (Webb et al. 2011). Also, we found no difference in sperm morphology between three Norwegian and one Czech population, which is in contrast with other studies showing geographic variation in sperm morphology (Lüpold et al. 2011; Schmoll and Kleven 2011). Furthermore, we found no evidence that nuclear divergence was related to mtDNA divergence. Even if the use of only two Z-linked loci limits the power to detect differences, it should be noted that we used sex-linked loci, which are more often differentiated between young species pairs relative to autosomes (Storchova et al. 2010; Hogner et al. 2012). Hence, if we assume that the divergence in mtDNA is a result of long periods of isolation with secondary contact, our data suggest that the redstart is undergoing speciation in reverse rather than early speciation.

Finally, deep mtDNA divergences may in theory evolve even in the absence of geographic isolation, provided that the effective population size is large enough (Webb et al. 2011). Such coexistence in one panmictic population would predict absence of geographic structure and reproductive barriers, similar mismatch distributions, and lack of divergence in the nuclear introns. We found support for most of these predictions, but the geographic structure in mtDNA and high variation in nuclear introns suggests that this hypothesis cannot fully explain the deep divergence in the common redstart. Possibly, the divergence arose in a period of isolation, for example, during one of the early Pleistocene glacial maxima, and continued to accumulate differences also after secondary contact had been achieved, due to large effective population sizes of both haplogroups. Alternatively, the two mtDNA lineages may have been subjected to differential selection pressures (e.g., local adaptation) that may have accelerated the divergence beyond neutral expectations, as recently suggested in a theoretical study by Irwin (2012).

We conclude that the deep, sympatric mtDNA lineages found in the common redstart do not represent cryptic species, nor are they likely to result from introgression from extant congenerics. Our data suggest that the divergence has evolved in isolated refugia, followed by secondary contact, or represent ancestral lineages that coexist in one panmictic population, or a combination of the two. Discriminating between these alternatives will

require deep genetic sampling combined with sophisticated multilocus, coalescence-based analyses. Sympatric mtDNA divergences are relatively rare in birds, but the fact that they occur argues against the use of threshold mtDNA divergences in species delineation.

Acknowledgments

We thank Peabody Museum of Natural History, Yale University (K. Zyskowski), Field Museum of Natural History Chicago (D. E. Willard), Natural History Museum of London (Tring, M. Adams) and Natural History Museum of Denmark, University of Copenhagen (J. Fjeldså and J. B. Kristensen), for providing skin samples. We thank E. Huhta, J. L. Copete, T. Øigarden, J. S. Hermansen, S. A. Sæther and F. Fossøy for providing samples, E. Rindal, G.-P. Sætre and V. Bakkestuen for computational and statistical assistance, A. Moksnes, E. Røskaft and B.G. Stokke for access to their study population in Røros, and L.E. Johannessen and J.A. Anmarkrud and E. Stensrud for help in the field. Funding was provided by the Natural History Museum, University of Oslo (PhD fellowship to SH).

Conflict of Interest

None declared.

References

ArisBrosou, S., and L. Excoffier. 1996. The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Mol. Biol. Evol. 13:494–504.

Avise, J. C. 2000. Phylogeography. Harvard Univ. press, Cambridge, Massachusetts London, England.

Avise, J. C., R. M. Ball, and J. Arnold. 1988. Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. Mol. Biol. Evol. 5:331–344.

Bachtrog, D., K. Thornton, A. Clark, and P. Andolfatto. 2006. Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. Evolution 60:292–302.

Barrowclough, G. F., J. G. Groth, K. J. Odom, and J. E. Lai. 2011. Phylogeography of the barred owl (*Strix varia*): species limits, multiple refugia, and range expansion. Auk 128:696–706.

Bensasson, D., D. X. Zhang, D. L. Hartl, and G. M. Hewitt. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol. Evol. 16:314–321.

Bertheau, C., H. Schuler, S. Krumbock, W. Arthofer, and C. Stauffer. 2011. Hit or miss in phylogeographic analyses: the case of the cryptic NUMTs. Mol. Ecol. Res. 11:1056–1059.

Birkhead, T. R., E. J. Pellatt, P. Brekke, R. Yeates, and H. Castillo-Juarez. 2005. Genetic effects on sperm design in the zebra finch. Nature 434:383–387.

Borge, T., M. T. Webster, G. Andersson, and G.-P. Sætre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. Genetics 171:1861–1873.

Bromham, L., and D. Penny. 2003. The modern molecular clock. Nat. Rev. Genet. 4:216–224.

Carr, C. M., S. M. Hardy, T. M. Brown, T. A. Macdonald, and P. D. N. Hebert. 2011. A Tri-Oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian Polychaetes. PLoS ONE 6:e22232.

Coyne, J. A., and A. H. Orr. 2004. Speciation. Sinauser Associates, Inc., Sunderland.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. 1:47–50.

Grant, P. R., and B. R. Grant. 1992. Hybridization of bird species. Science 256:193–197.

Grosch, K. 2004. Hybridization between redstart *Phoenicurus phoenicurus* and black redstart *P. ochruros*, and the effect on habitat exploitation. J. Avian Biol. 35:217–223.

Haine, E. R., J. Martin, and J. M. Cook. 2006. Deep mtDNA divergences indicate cryptic species in a fig-pollinating wasp. BMC Evol. Biol. 6:83.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12:101–109.

Hebert, P. D. N., S. Ratnasingham, and J. R. deWaard. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. Roy. Soc. Lond. B Biol. Sci. 270:S96–S99.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004a. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl Acad. Sci. USA 101:14812–14817.

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis. 2004b. Identification of birds through DNA barcodes. PLoS Biol. 2:1657–1663.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hogner, S., S. A. Sæther, T. Borge, T. Bruvik, A. Johnsen, and G.-P. Sætre. 2012. Increased divergence but reduced variation on the Z chromosome relative to autosomes in *Ficedula* flycatchers: differential introgression or the faster-Z effect? Ecol. Evol. 2:379–396.

del Hoyo, J., A. Elliot, and D. A. Christie. 2005. Handbook of the birds of the world. Lynx Edicions, Barcelona.

Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159.

Irwin, D. E. 2012. Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. Am. Nat. 180:35–49.

Johnsen, A., E. Rindal, P. Ericson, D. Zuccon, K. Kerr, M. Stoeckle, et al. 2010. DNA barcoding of Scandinavian birds reveals divergent lineages in trans-Atlantic species. J. Ornithol. 151:565–578.

Kerr, K. C. R., M. Y. Stoeckle, C. J. Dove, L. A. Weigt, C. M. Francis, and P. D. N. Hebert. 2007. Comprehensive DNA barcode coverage of North American birds. Mol. Ecol. Notes 7:535–543.

Kerr, K. C. R., S. M. Birks, M. V. Kalyakin, Y. A. Red'kin, E. A. Koblik, and P. D. N. Hebert. 2009a. Filling the gap – COI barcode resolution in eastern Palearctic birds. Front. Zool. 6:29.

Kerr, K. C. R., D. A. Lijtmaer, A. S. Barreira, P. D. N. Hebert, and P. L. Tubaro. 2009b. Probing evolutionary patterns in Neotropical Birds through DNA barcodes. PLoS ONE 4: e4379.

Kleven, O., T. Øigarden, B. E. Foyn, A. Moksnes, E. Røskaft, G. Rudolfsen, et al. 2007. Low frequency of extrapair paternity in the common redstart (Phoenicurus phoenicurus). J. Ornithol. 148:373–378.

Lande, R., S. Engen, and B.-E. Sæther. 2003. Stochastic Population Dynamics in Ecology and Conservation. Oxford Series in Ecology and Evolution, Oxford.

Laskemoen, T., O. Kleven, F. Fossøy, and J. T. Lifjeld. 2007. Intraspecific variation in sperm length in two passerine species, the bluethroat Luscinia svecica and the willow warbler Phylloscopus trochilus. Ornis Fennica 84:9.

Lavery, S., C. Moritz, and D. R. Fielder. 1996. Genetic patterns suggest exponential population growth in a declining species. Mol. Biol. Evol. 13:1106–1113.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lohman, D. J., K. K. Ingram, D. M. Prawiradilaga, K. Winker, F. H. Sheldon, R. G. Moyle, et al. 2010. Cryptic genetic diversity in "widespread" Southeast Asian bird species suggests that Philippine avian endemism is gravely underestimated. Biol. Conserv. 143:1885–1890.

Lüpold, S., D. F. Westneat, and T. R. Birkhead. 2011. Geographical variation in sperm morphology in the red-winged blackbird (*Agelaius phoeniceus*). Evol. Ecol. 25:373–390.

McCarthy, E. M. 2006. Handbook of Avian hybrids of the world. Oxford Univ. Press, New York.

Moritz, C., and C. Cicero. 2004. DNA barcoding: Promise and Pitfalls. PLoS Biol. 2:1529–1531.

Neigel, J. E., and J. C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation, in "Evolutionary Processes and Theory". Academic Press, New York.

Nosil, P., T. H. Vines, and D. J. Funk. 2005. Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution 59:705–719.

O'Neill, J. P., D. F. Lane, and L. N. Naka. 2011. A cryptic new species of thrush (*Turdudae: Turdus*) from Amazonia. Condor 113:869–880.

Päckert, M., J. Martens, D. T. Tietze, C. Dietzen, M. Wink, and L. Kvist. 2007. Calibration of a molecular clock in tits (*Paridae*) – Do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? Mol. Phylogenet. Evol. 44:1–14.

Pitnick, S., D. J. Hosken, and T. R. Birkhead. 2009. Sperm morphological diversity. Pp. 69–149 in T. R. Birkhead, D. J. Hosken and S. Pitnick, eds. Sperm biology: an evolutionary perspective. Academic Press, Oxford.

Price, T. 2008. Speciation in birds. Roberts and Company Publishers, Colorado.

Pulquério, M. J. F., and R. A. Nichols. 2007. Dates from the molecular clock: how wrong can we be? Trends Ecol. Evol. 22:180–184.

Quinn, T. W. 1992. The genetic legacy of Mother Goose-phylogeographic patterns of lesser snow goose (*Chen caerulescens caerulescens*) maternal lineages. Mol. Ecol. 1:105–117.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

Rogers, A. R., and H. Harpending. 1992. Population-growth makes waves in the distribution of pairwise genetic-differences. Mol. Biol. Evol. 9:552–569.

Sætre, G.-P., S. Riyahi, M. Aliabadian, J. S. Hermansen, S. Hogner, U. Olsson, et al. 2012. Single origin of human commensalism in the house sparrow. J. Evol. Biol. doi: 10.1111/j.1420-9101.2012.02470.x.

Schmoll, T., and O. Kleven. 2011. Sperm dimensions differ between two Coal Tit Periparus *ater* populations. J. Ornithol. 152:515–520.

Shapiro, L. H., R. A. Canterbury, D. M. Stover, R. C. Fleischer, and A. E. Burger. 2004. Reciprocal introgression between golden-winged warblers (*Vermivora chrysoptera*) and blue-winged warblers (*V. pinus*) in eastern north America. Auk 121:1019–1030.

Slatkin, M., and R. R. Hudson. 1991. Pairwise comparisons of the mitochondrial-DNA sequences in stable and exponentially growing populations. Genetics 129:555–562.

Sobel, J. M., G. F. Chen, L. R. Watt, and D. W. Schemske. 2010. The biology of speciation. Evolution 64:295–315.

Storchova, R., J. Reif, and M. W. Nachman. 2010. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. Evolution 64:456–471.

Tajima, F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA 4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596–1599.

Tarr, C. L. 1995. Primers for amplification and determination of mitochondrial control-region sequences in oscine passerines. Mol. Ecol. 4:527–529.

Tavares, E. S., and A. J. Baker. 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. BMC Evol. Biol. 8:81.

Taylor, S., A. Patirana, T. Birt, and V. Friesen. 2011. Cryptic introgression between murre sister species (Uria spp.) in the Pacific low Arctic: frequency, cause, and implications. Polar Biol. 35:1–10.

Toews, D. P. L., and A. Brelsford. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21:3907–3930.

Toews, D. P. L., and D. E. Irwin. 2008. Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. Mol. Ecol. 17:2691–2705.

Tominaga, K., K. Watanabe, R. Kakioka, S. Mori, and S.-R. Jeon. 2009. Two highly divergent mitochondrial DNA lineages within *Pseudogobio esocinus* populations in central Honshu, Japan. Ichthyol. Res. 56:195–199.

Triant, D. A., and L. D. Hayes. 2011. Molecular approaches in behavioural research: a cautionary note regarding mitochondrial transfers to the nucleus (numts). Anim. Behav. 82:601–606.

Wayne, R. K., A. Meyer, N. Lehman, B. Vanvalkenburg, P. W. Kat, T. K. Fuller, et al. 1990. Large sequence divergence among mitochondrial-DNA genotypes within populations of eastern african black-backed jackals. Proc. Natl Acad. Sci. USA 87:1772–1776.

Webb, W. C., J. M. Marzluff, and K. E. Omland. 2011. Random interbreeding between cryptic lineages of the common Raven: evidence for speciation in reverse. Mol. Ecol. 20:2390–2402.

Weir, B. S., and C. C. Cockerham. 1984. Estimating f-statistics for the analysis of population-structure. Evolution 38:1358–1370.

Weir, J. T., and D. Schluter. 2008. Calibrating the avian molecular clock. Mol. Ecol. 17:2321–2328.

Wolfson, A. 1952. The cloacal protuberance: a means for determining breeding condition in live male passerines. Bird Banding 23:7.

Xiao, J.-H., N.-X. Wang, R. W. Murphy, J. Cook, L.-Y. Jia, and D.-W. Huang. 2012. Wolbachia infection and dramatic intraspecific mitochondrial DNA divergence in a fig wasp. Evolution 66:1907–1916.

Zamudio, K. R., and W. K. Savage. 2003. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). Evolution 57:1631–1652.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Table S1. Detailes of all blood and high quality tissue samples.

Table S2. Details of the skin, feather samples and dried blood samples used.

Table S3. Details of the sequences downloaded from Genbank.

Figure S1. Neighbor-joining tree (K2P, 10,000 bootstrap replicates) based on COI (545bp) for contemporary common redstarts (N = 201), with black redstart as outgroup. The two common redstart haplogroups consist of 122 and 79 individuals, respectively. Only bootstrap values above 50% are shown. Blue = haplogroup 1 and red = haplogroup 2.

Figure S2. Neighbor-joining tree (K2P, 10,000 bootstrap replicates) based on a long fragment of COI (544bp) for seven *Phoenicurus* species. Only bootstrap values above 50% are shown. Blue = haplogroup 1 and red = haplogroup 2.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

SI Table 1. Detailes of all blood and high quality tissue samples.

Species	Sex	Date captured	Place captured	Collection/ sampled by	Journal nr	Ring nr	Source	Haplogroup
P. p. phoenicurus	Male	03.06.2007	Flåtevatnet Oppland,Norway	NHM, Oslo	20329	5E98118	Blood	2
P. p. phoenicurus	Male	14.06.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29527	7E44836	Blood	1
P. p. phoenicurus	Male	01.05.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29518	7E44805	Blood	2
P. p. phoenicurus	Male	21.05.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29519	7E44813	Blood	1
P. p. phoenicurus	Male	29.06.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29529	7E44849	Blood	1
P. p. phoenicurus	Male	21.05.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29521	7E44815	Blood	2
P. p. phoenicurus	Male	23.06.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	34196	7E44835	Blood	1
P. p. phoenicurus	Male	01.07.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29531	7E44851	Blood	2
P. p. phoenicurus	Female	24.06.2007	Flåtevatnet Oppland,Norway	NHM, Oslo	20330	5E98108	Blood	1
P. p. phoenicurus	Female	14.06.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29528	7E44837	Blood	1
P. p. phoenicurus	Female	21.05.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29520	7E44814	Blood	1
P. p. phoenicurus	Male	2003	Kamperud Østfold, Norway	NHM, Oslo	9811		Blood	2
P. p. phoenicurus	Male	18.06.2002	Røros, Norway	NHM, Oslo	14695	E770113	Blood	2

P. p. phoenicurus	Male	18.06.2002	Røros, Norway	NHM, Oslo	14704	E770122	Blood	2
P. p. phoenicurus	Male	17.06.2002	Røros, Norway	NHM, Oslo	14713	E770078	Blood	1
P. p. phoenicurus	Male	17.06.2002	Røros, Norway	NHM, Oslo	14722	E770087	Blood	2
P. p. phoenicurus	Male	08.06.2009	Røros, Norway	NHM, Oslo	37559	3E12274	Blood	2
P. p. phoenicurus	Male	08.06.2009	Røros, Norway	NHM, Oslo	37558	3E12273	Blood	1
P. p. phoenicurus	Male	09.06.2009	Røros, Norway	NHM, Oslo	37562	3E12278	Blood	1
P. p. phoenicurus	Male	09.06.2009	Røros, Norway	NHM, Oslo	37563	3E12279	Blood	2
P. p. phoenicurus	Male	09.06.2009	Røros, Norway	NHM, Oslo	37561	3E12277	Blood	1
P. p. phoenicurus	Male	08.06.2009	Røros, Norway	NHM, Oslo	37560	3E12275	Blood	1
P. p. phoenicurus	Unknown	08.06.2009	Røros, Norway	NHM, Oslo	37555	3E12270	Blood	2
P. p. phoenicurus	Unknown	08.06.2009	Røros, Norway	NHM, Oslo	37554	3E12269	Blood	1
P. p. phoenicurus	Unknown	08.06.2009	Røros, Norway	NHM, Oslo	37553	3E12268	Blood	1
P. p. phoenicurus	Male	14.06.2006	Røros, Norway	NHM, Oslo	25648		Blood	1
P. p. phoenicurus	Male	22.06.2006	Røros, Norway	NHM, Oslo	25657		Blood	1
P. p. phoenicurus	Male	23.06.2006	Røros, Norway	NHM, Oslo	25666		Blood	1
P. p. phoenicurus	Male	23.06.2006	Røros, Norway	NHM, Oslo	25675		Blood	2

P. p. phoenicurus	Male	23.06.2006	Røros, Norway	NHM, Oslo	25685		Blood	1
P. p. phoenicurus	Male	23.06.2006	Røros, Norway	NHM, Oslo	25693		Blood	2
P. p. phoenicurus	Female	2003	Kamperud Østfold, Norway	NHM, Oslo	9812		Blood	1
P. p. phoenicurus	Female	18.06.2002	Røros, Norway	NHM, Oslo	14696	E770112	Blood	1
P. p. phoenicurus	Female	18.06.2002	Røros, Norway	NHM, Oslo	14705	E770121	Blood	2
P. p. phoenicurus	Female	17.06.2002	Røros, Norway	NHM, Oslo	14714	E770077	Blood	2
P. p. phoenicurus	Female	17.06.2002	Røros, Norway	NHM, Oslo	14723	E770086	Blood	2
P. p. phoenicurus	Female	08.06.2009	Røros, Norway	NHM, Oslo	37556	3E12271	Blood	1
P. p. phoenicurus	Female	08.06.2009	Røros, Norway	NHM, Oslo	37557	3E12272	Blood	1
P. p. phoenicurus	Female	09.06.2009	Røros, Norway	NHM, Oslo	37564	3E12280	Blood	2
P. p. phoenicurus	Female	14.06.2006	Røros, Norway	NHM, Oslo	25649		Blood	2
P. p. phoenicurus	Female	22.06.2006	Røros, Norway	NHM, Oslo	25658		Blood	1
P. p. phoenicurus	Female	23.06.2006	Røros, Norway	NHM, Oslo	25667		Blood	2
P. p. phoenicurus	Female	23.06.2006	Røros, Norway	NHM, Oslo	25676		Blood	2
P. p. phoenicurus	Female	23.06.2006	Røros, Norway	NHM, Oslo	25686		Blood	2
P. p. phoenicurus	Female	23.06.2006	Røros, Norway	NHM, Oslo	25694		Blood	2

P. p. phoenicurus	Male	12.06.2009	Trysil, Norway	NHM, Oslo	29036	3E12874	Blood	2
P. p. phoenicurus	Male	12.06.2009	Trysil, Norway	NHM, Oslo	29044	3E12882	Blood	1
P. p. phoenicurus	Male	12.06.2009	Trysil, Norway	NHM, Oslo	29045	3E12883	Blood	2
P. p. phoenicurus	Female	11.06.2009	Trysil, Norway	NHM, Oslo	29037	3E12875	Blood	2
P. p. phoenicurus	Female	12.06.2009	Trysil, Norway	NHM, Oslo	29043	3E12881	Blood	1
P. p. phoenicurus	Female	10.06.2009	Trysil, Norway	NHM, Oslo	29030	3E12868	Blood	2
P. p. phoenicurus	Male	19.06.2002	Trysil, Norway	NHM, Oslo	14731	2E95422	Blood	1
P. p. phoenicurus	Male	20.06.2002	Trysil, Norway	NHM, Oslo	14740	2E95455	Blood	2
P. p. phoenicurus	Male	21.06.2002	Trysil, Norway	NHM, Oslo	14748	2E95230	Blood	2
P. p. phoenicurus	Male	21.06.2002	Trysil, Norway	NHM, Oslo	14757	2E95463	Blood	2
P. p. phoenicurus	Male	21.06.2002	Trysil, Norway	NHM, Oslo	14790	2E95460	Blood	2
P. p. phoenicurus	Female	19.06.2002	Trysil, Norway	NHM, Oslo	14732	2E95423	Blood	1
P. p. phoenicurus	Female	19.06.2002	Trysil, Norway	NHM, Oslo	14741	2E95430	Blood	2
P. p. phoenicurus	Female	21.06.2002	Trysil, Norway	NHM, Oslo	14749	2E95229	Blood	2
P. p. phoenicurus	Female	21.06.2002	Trysil, Norway	NHM, Oslo	14758	2E95462	Blood	2
P. p. phoenicurus	Female	21.06.2002	Trysil, Norway	NHM, Oslo	14815	2E95461	Blood	1

P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29038	3E12876	Blood	2
P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29035	3E12873	Blood	1
P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29034	3E12872	Blood	2
P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29033	3E12871	Blood	1
P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29032	3E12870	Blood	2
P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29031	3E12869	Blood	2
P. p. phoenicurus	Male	10.06.2009	Trysil, Norway	NHM, Oslo	37565	7E44826	Blood	1
P. p. phoenicurus	Male	11.06.2009	Trysil, Norway	NHM, Oslo	29039	3E12877	Blood	2
P. p. phoenicurus	Female	11.06.2006	Trysil, Norway	NHM, Oslo	29040	3E12878	Blood	2
P. p. phoenicurus	Male	11.06.2006	Trysil, Norway	NHM, Oslo	29041	3E12879	Blood	1
P. p. phoenicurus	Male	11.06.2006	Trysil, Norway	NHM, Oslo	29042	3E12880	Blood	1
P. p. phoenicurus	Male	12.06.2006	Trysil, Norway	NHM, Oslo	29046	3E12884	Blood	1
P. p. phoenicurus	Male	14.06.2006	Trysil, Norway	NHM, Oslo	25505	6E69119	Blood	2
P. p. phoenicurus	Male	22.06.2006	Trysil, Norway	NHM, Oslo	25521	6E69150	Blood	2
P. p. phoenicurus	Male	16.06.2006	Trysil, Norway	NHM, Oslo	25530	6E69147	Blood	1
P. p. phoenicurus	Male	14.06.2006	Trysil, Norway	NHM, Oslo	25538	6E69121	Blood	2

P. p. phoenicurus	Male	22.06.2006	Trysil, Norway	NHM, Oslo	25543	6E69146	Blood	2
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25551	6E69138	Blood	1
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25568	6E69129	Blood	2
P. p. phoenicurus	Male	14.06.2006	Trysil, Norway	NHM, Oslo	25591	6E69124	Blood	2
P. p. phoenicurus	Male	14.06.2006	Trysil, Norway	NHM, Oslo	25606	6E69128	Blood	1
P. p. phoenicurus	Male	16.06.2006	Trysil, Norway	NHM, Oslo	25615	6E69143	Blood	2
P. p. phoenicurus	Male	17.06.2006	Trysil, Norway	NHM, Oslo	25623	6E69144	Blood	1
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25640	6E69132	Blood	2
P. p. phoenicurus	Male	19.06.2002	Trysil, Norway	NHM, Oslo	14765	2E95453	Blood	1
P. p. phoenicurus	Male	20.06.2002	Trysil, Norway	NHM, Oslo	14773	2E95458	Blood	1
P. p. phoenicurus	Male	20.06.2002	Trysil, Norway	NHM, Oslo	14781	2E95456	Blood	2
P. p. phoenicurus	Male	20.06.2002	Trysil, Norway	NHM, Oslo	14793	2E95454	Blood	2
P. p. phoenicurus	Male	21.06.2002	Trysil, Norway	NHM, Oslo	14789	2E95459	Blood	1
P. p. phoenicurus	Male	16.06.2006	Trysil, Norway	NHM, Oslo	25513	6E69140	Blood	2
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25561	6E69137	Blood	1
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25577	6E69136	Blood	1

P. p. phoenicurus	Male	19.06.2006	Trysil, Norway	NHM, Oslo	25584	6E69145	Blood	2
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25598	6E69135	Blood	1
P. p. phoenicurus	Male	15.06.2006	Trysil, Norway	NHM, Oslo	25632	6E69133	Blood	1
P. p. phoenicurus	Female	14.06.2006	Trysil, Norway	NHM, Oslo	25506	6E69118	Blood	2
P. p. phoenicurus	Female	22.06.2006	Trysil, Norway	NHM, Oslo	25522	6E69149	Blood	2
P. p. phoenicurus	Female	16.06.2006	Trysil, Norway	NHM, Oslo	25531	6E69141	Blood	1
P. p. phoenicurus	Female	14.06.2006	Trysil, Norway	NHM, Oslo	25539	6E69120	Blood	1
P. p. phoenicurus	Female	16.06.2006	Trysil, Norway	NHM, Oslo	25544	6E69142	Blood	2
P. p. phoenicurus	Female	14.06.2006	Trysil, Norway	NHM, Oslo	25552	6E69122	Blood	2
P. p. phoenicurus	Female	14.06.2006	Trysil, Norway	NHM, Oslo	25569	6E69123	Blood	2
P. p. phoenicurus	Female	14.06.2006	Trysil, Norway	NHM, Oslo	25592	6E69126	Blood	2
P. p. phoenicurus	Female	14.06.2006	Trysil, Norway	NHM, Oslo	25607	6E69127	Blood	2
P. p. phoenicurus	Female	15.06.2006	Trysil, Norway	NHM, Oslo	25616	6E69130	Blood	2
P. p. phoenicurus	Female	15.06.2006	Trysil, Norway	NHM, Oslo	25624	6E69134	Blood	1
P. p. phoenicurus	Female	15.06.2006	Trysil, Norway	NHM, Oslo	25641	6E69131	Blood	2
P. p. phoenicurus	Chick	2002	Trysil, Norway	NHM, Oslo	14733		Blood	1

P. p. phoenicurus	Chick	2002	Trysil, Norway	NHM, Oslo	14742		Blood	1
P. p. phoenicurus	Chick	2002	Trysil, Norway	NHM, Oslo	14774		Blood	2
P. p. phoenicurus	Chick	2002	Trysil, Norway	NHM, Oslo	14782		Blood	1
P. p. phoenicurus	Chick	2002	Trysil, Norway	NHM, Oslo	14808		Blood	2
P. p. phoenicurus	Chick	2006	Trysil, Norway	NHM, Oslo	25515		Blood	2
P. p. phoenicurus	Chick	2006	Trysil, Norway	NHM, Oslo	25563		Blood	2
P. p. phoenicurus	Chick	2006	Trysil, Norway	NHM, Oslo	25578		Blood	1
P. p. phoenicurus	Chick	2006	Trysil, Norway	NHM, Oslo	25585		Blood	1
P. p. phoenicurus	Chick	2006	Trysil, Norway	NHM, Oslo	25599		Blood	2
P. p. phoenicurus	Chick	2006	Trysil, Norway	NHM, Oslo	25633		Blood	1
P. p. phoenicurus	Male	01.07.2009	Flåtevatnet Oppland,Norway	NHM, Oslo	29532	7E44852	Blood	2
P. p. phoenicurus	Male	22.06.2006	Trysil, Norway	NHM, Oslo	25521	6E69150	Blood	2
P. p. phoenicurus	Male	10.05.2007	Norway	NHM, Oslo	37552	3E12142	Blood	2
P. p. phoenicurus	Male	16.06.2008	Sør-Audal, Oppland,Norway	NHM, Oslo	24190	7E43078	Blood	2
P. p. phoenicurus	Male	16.06.2008	Sør-Audal, Oppland,Norway	NHM, Oslo	24198	7E43079	Blood	2
P. p. phoenicurus	Male	16.06.2008	Sør-Audal, Oppland,Norway	NHM, Oslo	24205	7E43081	Blood	1

P. p. phoenicurus	Male	23.05.2008	Ifrane, Morocco	NHM, Oslo	24303		Blood	1
P. p. phoenicurus	Male	02.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37494	TE95873	Blood	1
P. p. phoenicurus	Male	03.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37495	TE95875	Blood	1
P. p. phoenicurus	Male	03.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37496	TE95876	Blood	1
P. p. phoenicurus	Male	04.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37471	TE95879	Blood	1
P. p. phoenicurus	Male	04.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37497	TE95880	Blood	1
P. p. phoenicurus	Male	05.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37453	TE95881	Blood	2
P. p. phoenicurus	Male	21.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37461	TE95889	Blood	1
P. p. phoenicurus	Male	04.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37469	TE95877	Blood	2
P. p. phoenicurus	Male	02.07.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	40509	TK22632	Blood	1
P. p. phoenicurus	Male	29.06.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	40508	TK22631	Blood	1
P. p. phoenicurus	Male	23.06.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	40505	TK22627	Blood	1
P. p. phoenicurus	Female	04.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37472	TE95878	Blood	1
P. p. phoenicurus	Male	03.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37452	TE95874	Blood	2
P. p. phoenicurus	Female	21.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37460	TE95888	Blood	1
P. p. phoenicurus	Female	27.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37470	TE95897	Blood	1

P. p. phoenicurus	Female	25.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37487	TE95890	Blood	1
P. p. phoenicurus	Female	24.06.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	40507	TK22630	Blood	1
P. p. phoenicurus	Female	24.06.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	40506	TK22629	Blood	1
P. p. phoenicurus	Female	23.06.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	40504	TK22626	Blood	1
P. p. phoenicurus	Female	21.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37453	TE95887	Blood	1
P. p. phoenicurus	Female	05.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37485	TE95883	Blood	1
P. p. phoenicurus	Female	21.05.2010	Hradec Kràlové, Czech Republic	NHM, Oslo	37460	TE95888	Blood	1
P. p. phoenicurus	Female	10.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34184	ED12453	Blood	2
P. p. phoenicurus	Female	16.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34187	E709938	Blood	2
P. p. phoenicurus	Male	16.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34188	E709939	Blood	2
P. p. phoenicurus	Female	16.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34185	E709936	Blood	2
P. p. phoenicurus	Male	08.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34181	E709934	Blood	2
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37380	ED12420	Blood	1
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37360	ED12408	Blood	2
P. p. phoenicurus	Female	15.06.2010	Røros, Norway	NHM, Oslo	37399	3E12563	Blood	2
P. p. phoenicurus	Female	19.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34189	E709946	Blood	2

P. p. phoenicurus	Male	10.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34190	E709947	Blood	1
P. p. phoenicurus	Male	10.06.2010	Trysil, Norway	NHM, Oslo	37346	ED12403	Blood	2
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37370	ED12414	Blood	2
P. p. phoenicurus	Male	10.06.2010	Trysil, Norway	NHM, Oslo	37368	ED12412	Blood	2
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37379	ED12419	Blood	2
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37369	ED12413	Blood	1
P. p. phoenicurus	Female	23.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34195	E709993	Blood	2
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37342	ED12401	Blood	2
P. p. phoenicurus	Male	10.10.2010	Trysil, Norway	NHM, Oslo	37343	ED12402	Blood	2
P. p. phoenicurus	Male	08.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34180	E709933	Blood	2
P. p. phoenicurus	Male	09.05.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34162	E709915	Blood	1
P. p. phoenicurus	Male	07.05.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34176	E709930	Blood	1
P. p. phoenicurus	Male	07.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34178	E709931	Blood	1
P. p. phoenicurus	Female	23.05.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34193	E709973	Blood	1
P. p. phoenicurus	Female	30.05.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	20354	5E98117	Blood	1
P. p. phoenicurus	Female	08.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34179	E709932	Blood	1

P. p. phoenicurus	Female	16.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34186	E709937	Blood	1
P. p. phoenicurus	Chick	23.05.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34194	E709987	Blood	2
P. p. phoenicurus	Female	15.06.2010	Røros, Norway	NHM, Oslo	37419	7E44014	Blood	1
P. p. phoenicurus	Female	15.06.2010	Røros, Norway	NHM, Oslo	37420	7E44015	Blood	2
P. p. phoenicurus	Female	15.06.2010	Røros, Norway	NHM, Oslo	37421	7E44016	Blood	1
P. p. phoenicurus	Female	15.06.2010	Røros, Norway	NHM, Oslo	37422	7E44017	Blood	1
P. p. phoenicurus	Female	15.06.2010	Røros, Norway	NHM, Oslo	37423	7E44018	Blood	1
P. p. phoenicurus	Male	15.06.2010	Røros, Norway	NHM, Oslo	37397	3E12562	Blood	1
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37374	ED12415	Blood	1
P. p. phoenicurus	Male	14.06.2010	Røros, Norway	NHM, Oslo	37384	3E12558	Blood	1
P. p. phoenicurus	Male	16.06.2010	Røros, Norway	NHM, Oslo	37407	3E12566	Blood	1
P. p. phoenicurus	Male	15.06.2010	Røros, Norway	NHM, Oslo	37395	3E12561	Blood	1
P. p. phoenicurus	Female	14.06.2010	Røros, Norway	NHM, Oslo	37383	3E12557	Blood	2
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37376	ED12417	Blood	2
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37379	ED12419	Blood	2

P. p. phoenicurus	Male	10.06.2010	Trysil, Norway	NHM, Oslo	37364	ED12411	Blood	1
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37375	ED12416	Blood	1
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37347	ED12404	Blood	1
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37354	ED12407	Blood	1
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37363	ED12410	Blood	1
P. p. phoenicurus	Male	27.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34198	EC33312	Blood	2
P. p. phoenicurus	Egg	22.06.2010	Trysil, Norway	NHM, Oslo	40529		Egg	1
P. p. phoenicurus	Female	27.06.2010	Flåtevatnet Oppland,Norway	NHM, Oslo	34197	EC33311	Blood	1
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37377	ED12418	Blood	1
P. p. phoenicurus	Male	11.06.2010	Trysil, Norway	NHM, Oslo	37381	ED12421	Blood	2
P. p. phoenicurus	Male	22.06.2010	Trysil, Norway	NHM, Oslo	37425	ED12423	Blood	1
P. p. phoenicurus	Male	22.06.2010	Trysil, Norway	NHM, Oslo	37426	ED12424	Blood	1
P. p. phoenicurus	Male	22.06.2010	Trysil, Norway	NHM, Oslo	37428	ED12426	Blood	1
P. p. phoenicurus	Female	17.06.2002	Røros, Norway	NHM, Oslo	14864	E770072	Blood	2
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	14807	ED12452	Blood	2

P. p. phoenicurus	Female	17.06.2002	Røros, Norway	NHM, Oslo	14865	E770073	Blood	2
P. p. phoenicurus	Female	18.06.2002	Røros, Norway	NHM, Oslo	14840	E770103	Blood	2
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	37349	ED12405	Blood	2
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	34182	ED12451	Blood	2
P. p. phoenicurus	Female	18.06.2002	Elverum, Norway	NHM, Oslo	14791	1E98294	Blood	2
P. p. phoenicurus	Female	17.06.2002	Røros, Norway	NHM, Oslo	14871	E770076	Blood	2
P. p. phoenicurus	Female	19.06.2002	Trysil, Norway	NHM, Oslo	14800	2E95438	Blood	1
P. p. phoenicurus	Female	18.06.2002	Røros, Norway	NHM, Oslo	14848	E770059	Blood	1
P. p. phoenicurus	Female	10.06.2010	Trysil, Norway	NHM, Oslo	14792	2E95451	Blood	1
P. p. phoenicurus	Female	17.06.2002	Røros, Norway	NHM, Oslo	14856	E770071	Blood	1
P. p. phoenicurus	Female	18.06.2002	Røros, Norway	NHM, Oslo	14839	E770095	Blood	1
P. p. phoenicurus	Female	18.06.2002	Røros, Norway	NHM, Oslo	14841	E770120	Blood	1
P. p. phoenicurus	Female	22.06.2010	Trysil, Norway	NHM, Oslo	37427	ED12425	Blood	1
P. p. phoenicurus	Male	03.06.2010	Dovre, Norway	NHM, Oslo	40512	9E91349	Blood	2
P. p. phoenicurus	Unknown	02.08.2010	Dovre, Norway	NHM, Oslo	40517	9E91513	Blood	1
P. p. phoenicurus	Female	09.08.2010	Dovre, Norway	NHM, Oslo	40521	9E91593	Blood	1

P. p. phoenicurus	Female	24.06.2010	Dovre, Norway	NHM, Oslo	40516	9E91471	Blood	2
P. p. phoenicurus	Unknown	05.08.2010	Dovre, Norway	NHM, Oslo	40520	9E91547	Blood	1
P. p. samamisicus	Chick	06.05.2010	Antalya, Turkey	NHM, Oslo	37498		Feather	1
P. p. samamisicus	Chick	06.05.2010	Antalya, Turkey	NHM, Oslo	37499		Feather	1
P. p. samamisicus	Chick	16.05.2010	Antalya, Turkey	NHM, Oslo	37514		Feather	1
P. p. samamisicus	Female	16.05.2010	Antalya, Turkey	NHM, Oslo	37520		Tissue	1
P. p. samamisicus	Male	16.05.2010	Antalya, Turkey	NHM, Oslo	37521		Tissue	1
P. p. samamisicus	Female	17.05.2010	Antalya, Turkey	NHM, Oslo	37529		Tissue	1
P. p. samamisicus	Male	17.05.2010	Antalya, Turkey	NHM, Oslo	37530		Tissue	1
P. p. samamisicus	Chick	06.06.2010	Antalya, Turkey	NHM, Oslo	37537		Feather	1
P. p. samamisicus	Male	24.05.2010	Antalya, Turkey	NHM, Oslo	37542		Tissue	1
P. p. samamisicus	Chick	06.06.2010	Antalya, Turkey	NHM, Oslo	37536		Feather	1
P. p. samamisicus	Male	06.05.2010	Antalya, Turkey	NHM, Oslo	37511		Blood	1
P. p. samamisicus	Female	06.05.2010	Antalya, Turkey	NHM, Oslo	37510		Blood	1
P. p. samamisicus	Female	06.05.2010	Antalya, Turkey	NHM, Oslo	37504		Blood	1
P. p. samamisicus	Male	06.05.2010	Antalya, Turkey	NHM, Oslo	37512		Blood	1

P. p. samamisicus	Male	06.05.2010	Antalya, Turkey	NHM, Oslo	37513	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38644	Blood	1
P. p. samamisicus	Male	03.06.2010	Chika, Jahan Nama, Iran	NHM, Oslo	38643	Blood	1
P. p. samamisicus	Male	02.06.2010	Chika, Jahan Nama, Iran	NHM, Oslo	38642	Blood	1
P. p. samamisicus	Female	01.06.2010	Chika, Jahan Nama, Iran	NHM, Oslo	38640	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38645	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38646	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38647	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38648	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38649	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38650	Blood	1
P. p. samamisicus	Female	08.06.2010	Mazichal, Iran	NHM, Oslo	38654	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38658	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38653	Blood	1
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38651	Blood	2
P. p. samamisicus	Female	08.06.2010	Mazichal, Iran	NHM, Oslo	38652	Blood	2

P. p. samamisicus	Female	08.06.2010	Mazichal, Iran	NHM, Oslo	38656	Blood	2
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38655	Blood	2
P. p. samamisicus	Male	08.06.2010	Mazichal, Iran	NHM, Oslo	38657	Blood	2
P. p. phoenicurus	Male	12.05.2008	Valjevo, Serbia	YPM, New Haven	84332	Tissue	1
P. p. phoenicurus	Male	12.05.2008	Valjevo, Serbia	YPM, New Haven	84337	Tissue	1
P. p. phoenicurus	Female	12.05.2008	Valjevo, Serbia	YPM, New Haven	84342	Tissue	2
P. p. phoenicurus	Male	09.05.2008	Valjevo, Serbia	YPM, New Haven	84284	Tissue	2
P. p. phoenicurus	Female	09.05.2008	Valjevo, Serbia	YPM, New Haven	84301	Tissue	2
P. p. phoenicurus	Unknown	09.05.2008	Valjevo, Serbia	YPM, New Haven	84302	Tissue	1
P. p. phoenicurus	Unknown	09.05.2008	Valjevo, Serbia	YPM, New Haven	84303	Tissue	1
P. p. phoenicurus	Female	15.05.2005	Krasnodarskiy Kray, Russia	YPM, New Haven	140371	Tissue	1
P. p. phoenicurus	Female	15.05.2005	Krasnodarskiy Kray, Russia	YPM, New Haven	140372	Tissue	1
P. p. phoenicurus	Female	15.05.2005	Krasnodarskiy Kray, Russia	YPM, New Haven	140373	Tissue	1
P. p. phoenicurus	Female	21.05.2005	Krasnodarskiy Kray, Russia	YPM, New Haven	140803	Tissue	1
P. p. phoenicurus	Male	21.05.2005	Krasnodarskiy Kray, Russia	YPM, New Haven	140804	Tissue	2
P. p. phoenicurus	Male	28.05.2004	Krasnodarskiy Kray, Russia	YPM, New Haven	140824	Tissue	1

P. p. phoenicurus	Female	28.05.2004	Krasnodarskiy Kray, Russia	YPM, New Haven	140827		Tissue	1
P. p. phoenicurus	Male	29.05.2004	Krasnodarskiy Kray, Russia	YPM, New Haven	140730		Tissue	1
P. p. phoenicurus	Male	30.05.2004	Krasnodarskiy Kray, Russia	YPM, New Haven	140837		Tissue	1
P. p. phoenicurus	Male	19.06.2006	Krasnodarskiy Kray, Russia	YPM, New Haven	101876		Tissue	1
P. p. phoenicurus	Female	23.06.2006	Krasnodarskiy Kray, Russia	YPM, New Haven	101813		Tissue	1
P. p. phoenicurus	Male	23.06.2006	Krasnodarskiy Kray, Russia	YPM, New Haven	101905		Tissue	1
P. p. phoenicurus	Female	14.06.2004	Republic of Adygea, Russia	YPM, New Haven	141204		Tissue	1
P. p. phoenicurus	Male	15.06.2004	Republic of Adygea, Russia	YPM, New Haven	140880		Tissue	2
P. p. phoenicurus	Male	15.06.2004	Republic of Adygea, Russia	YPM, New Haven	141206		Tissue	1
P. p. phoenicurus	Male	15.06.2004	Republic of Adygea, Russia	YPM, New Haven	141207		Tissue	2
P. p. phoenicurus	Male	05.05.2012	Viladrau, Spain	NHM, Oslo	38969	HH2992	Blood	2
P. p. phoenicurus	Male	05.05.2012	Viladrau, Spain	NHM, Oslo	68970	HH2995	Blood	1
P. p. phoenicurus	Male	05.05.2012	Viladrau, Spain	NHM, Oslo	38971	HH3000	Blood	2
P. p. phoenicurus	Male	19.05.2012	Viladrau, Spain	NHM, Oslo	38972	KH7974	Blood	1
P. p. phoenicurus	Female	02.06.2012	Viladrau, Spain	NHM, Oslo	38973	KH7975	Blood	2
P. p. phoenicurus	Male	02.06.2012	Viladrau, Spain	NHM, Oslo	38974	KH7976	Blood	2

P. p. samamisicus	Male	22.04.2008	Eilat, Israel	NHM, Oslo	23855		Blood	1
P. p. samamisicus	Female	17.09.2007	Eilat, Israel	NHM, Oslo	22681		Blood	2
P. p. samamisicus	Male	22.04.2008	Eilat, Israel	NHM, Oslo	23854		Blood	1
P. p. samamisicus	Male	21.04.2008	Eilat, Israel	NHM, Oslo	23850		Blood	1
P. p. samamisicus	Male	18.04.2008	Eilat, Israel	NHM, Oslo	23852		Blood	1
P. p. samamisicus	Male	16.10.2007	Eilat, Israel	NHM, Oslo	22689		Blood	1
P. p. samamisicus	Female	09.10.2007	Eilat, Israel	NHM, Oslo	22683		Blood	1
P. p. samamisicus	Female	12.10.2007	Eilat, Israel	NHM, Oslo	22687		Blood	1
P. p. samamisicus	Unknown	27.09.2007	Eilat, Israel	NHM, Oslo	22682		Blood	1
P. p. samamisicus	Male	18.04.2008	Eilat, Israel	NHM, Oslo	23851		Blood	1
P. p. samamisicus	Male	12.10.2007	Eilat, Israel	NHM, Oslo	22686		Blood	1
P. p. samamisicus	Male	10.10.2007	Eilat, Israel	NHM, Oslo	22685		Blood	1
P. p. samamisicus	Male	10.10.2007	Eilat, Israel	NHM, Oslo	22684		Blood	1
P. p. samamisicus	Female	16.10.2007	Eilat, Israel	NHM, Oslo	22688		Blood	1
P. p. phoenicurus	Male	02.06.2012	Viladrau, Spain	NHM, Oslo	38975	KH7977	Blood	1
P. p. phoenicurus	Male	16.06.2012	Viladrau, Spain	NHM, Oslo	38976	KH7978	Blood	2

P. moussieri	Unknown	18.11.2004	Tunisia	NHM, Copenhagen	141957	Tissue
P.frontalis	Unknown	20.06.1990	Sichuan, China	NHM, Copenhagen	117098	Tissue
P. erythronotus	Unknown	21.07.2008	Tien-Shan mtn, Kyrgyzstan	NHM, Copenhagen	141249	Tissue
P. erythronotus	Male	20.07.2008	Tien-Shan mtn, Kyrgyzstan	NHM, Copenhagen	141218	Tissue

NHM, Oslo = Natural History Museum, Oslo, Norway; YPM, New Haven = Yale Peabody Museum of Natural History, New Haven, USA; NHM, Copenhagen= Natural History Museum, Copenhagen, Denmark.

SI Table 2. Details of the skin, feather samples and dried blood samples used

Species	Sex	Date captured	Place captured	Collection/ sampled by	Journal nr	Ring nr	Source	Haplogroup
P. phoenicurus	Male	01.07.1908	Surnadal, Norway	NHM, Oslo	1505		Skin	1
P. phoenicurus	Male	15.07.1886	Vestre Aker, Norway	NHM, Oslo	1499		Skin	2
P. phoenicurus	Male	20.07.1866	Finnmark, Norway	NHM, Oslo	1495		Skin	2
P. phoenicurus	Female	26.09.1908	Jæren, Norway	NHM, Oslo	1506		Skin	1
P. phoenicurus	Male	18.06.1966	Øvre Pasvik, Norway	NHM, Oslo	8868		Skin	1
P. phoenicurus	Unknown	29.06.1886	Porsgrunn, Norway	NHM, Oslo	1883		Skin	1
P. phoenicurus	Male	18.06.1966	Øvre Pasvik, Norway	NHM, Oslo	8869		Skin	2
P. phoenicurus	Male	18.06.1966	Øvre Pasvik, Norway	NHM, Oslo	8870		Skin	2
P. phoenicurus	Unknown	22.07.1882	Drammen, Norway	NHM, Oslo	1497		Skin	1
P. phoenicurus	Unknown	10.07.1904	Gausdal, Norway	NHM, Oslo	1504		Skin	1
P. phoenicurus	Female	27.06.1992	Meltaus area, Finland	Esa Huhta		084393	Feather	1
P. phoenicurus	Female	30.06.1994	Meltaus area, Finland	Esa Huhta		367928	Feather	2
P. phoenicurus	Female	26.06.1992	Meltaus area, Finland	Esa Huhta		084390	Feather	2

P. phoenicurus	Female	07.06.1993	Meltaus area, Finland	Esa Huhta	264871	Feather	1
P. phoenicurus	Female	27.06.1992	Meltaus area, Finland	Esa Huhta	084392	Feather	1
P. phoenicurus	Female	04.06.1992	Meltaus area, Finland	Esa Huhta	264785	Feather	2
P. phoenicurus	Female	25.06.1992	Meltaus area, Finland	Esa Huhta	264759	Feather	1
P. phoenicurus	Female	28.06.1993	Meltaus area, Finland	Esa Huhta	367883	Feather	2
P. phoenicurus	Female	16.06.1993	Meltaus area, Finland	Esa Huhta	367816	Feather	1
P. p. phoenicurus	Female	22.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r842703	Blood	1
P. p. phoenicurus	Female	22.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r521699J	Blood	2
P. p. phoenicurus	Female	19.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r521691J	Blood	2
P. p. phoenicurus	Female	30.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r842735J	Blood	2
P. p. phoenicurus	Female	22.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r521700J	Blood	2
P. p. phoenicurus	Female	09.07.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r842736J	Blood	1
P. p. phoenicurus	Female	22.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy	r842701J	Blood	1

P. p. phoenicurus	Female	22.06.2006	North Karelia, Joensuu, Finland	Frode Fossøy		r521692J	Blood	1
P. phoenicurus	Chick	1999	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37992		Dried blood	2
P. phoenicurus	Chick	1999	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37993		Dried blood	1
P. phoenicurus	Chick	1998	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37994		Dried blood	1
P. phoenicurus	Chick	1998	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37995		Dried blood	2
P. phoenicurus	Chick	1998	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37996		Dried blood	1
P. phoenicurus	Chick	1998	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37997		Dried blood	2
P. phoenicurus	Chick	1998	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37998		Dried blood	1
P. phoenicurus	Chick	1998	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	37999		Dried blood	1
P. phoenicurus	Chick	1999	Hradec Kràlové, Czech Republic	Jiri Porkert/ NHM Oslo	38001		Dried blood	1
P. phoenicurus	Male	28.05.1929	Sør-Varanger, Norway	NHM, Copenhagen	10.725		Skin	2
P. phoenicurus	Male	August, 1930	Lolland, Denmark	NHM, Copenhagen	19.839		Skin	2
P. phoenicurus	Male	15.06.1890	Denmark	NHM, Copenhagen	19.845		Skin	1
P. phoenicurus	Male	25.06.1936	Fyrskib, Denmark	NHM, Copenhagen	19.867		Skin	1
P. phoenicurus	Female	01.09.1905	Fornas fyr, Denmark	NHM, Copenhagen	19.869		Skin	1

P. phoenicurus	Female	01.09.1894	Lodbjerg fyr, Denmark	NHM, Copenhagen	19.787	Skin	1
P. phoenicurus	Female	28.06.1959	Copenhagen, Denmark	NHM, Copenhagen	70.991	Skin	1
P. phoenicurus	Unknown	22.06.1895	Klampenborg, Denmark	NHM, Copenhagen	19.877	Skin	1
P. phoenicurus	Male	07.08.1977	Sjælland, Denmark	NHM, Copenhagen	67.349	Skin	2
P. phoenicurus	Male	09.07.1915	Lule, Lappmark, Sweden	NHM, Copenhagen	14.542	Skin	2
P. phoenicurus	Male	16.07.1915	Lule, Lappmark, Sweden	NHM, Copenhagen	4.740	Skin	1
P. phoenicurus	Male	26.06.1929	Øvre Passvik, Norway	NHM, Copenhagen	13.266	Skin	2
P. phoenicurus	Male	22.09.1909	Nakkehoved fyr, Denmark	NHM, Copenhagen	19.795	Skin	2
P. phoenicurus	Male	10.05.1946	Copenhagen, Denmark	NHM, Copenhagen	70.989	Skin	1
P. phoenicurus	Male	20.05	Schleswig Holstein, Germany	HZM, London	HZM.83.4468 /199	Skin	1
P. phoenicurus	Male	03.07.1942	Westerlane, Kent, England	HZM, London	HZM.76.4461	Skin	1
P. phoenicurus	Male	01.07.1937	Bern, Switzerland	HZM, London	HZM.75.4460	Skin	1
P. phoenicurus	Female	03.07.1942	Sevenoaks, Kent, England	HZM, London	HZM.61.4446	Skin	1
P. phoenicurus	Male	17.06.1939	Bern, Switzerland	HZM, London	HZM.59.4444	Skin	1
P. phoenicurus	Male	12.05.1951	Biel, Bern, Switzerland	HZM, London	HZM.46.4431	Skin	1

P. phoenicurus	Male	07.08.1919	Kincraig, Scotland	HZM, London	HZM.33.4418	Skin	1
P. phoenicurus	Female	07.08.1919	Kincraig, Scotland	HZM, London	HZM.44.4426	Skin	1
P. phoenicurus	Male	21.07.1952	Saxony, Germany	HZM, London	HZM.39.4424	Skin	1
P. phoenicurus	Male	11.05.1937	Blois, France	HZM, London	HZM.110.18898	Skin	1
P. phoenicurus	Male	10.07.1934	Bern, Switzerland	HZM, London	HZM.17.4402	Skin	1
P. phoenicurus	Male	18.07.1932	Bern, Switzerland	HZM, London	HZM.11.4396	Skin	2
P. phoenicurus	Male	13.06.1939	Bern, Switzerland	HZM, London	HZM.13.4398	Skin	1
P. phoenicurus	Male	24.07.1953	Bern, Switzerland	HZM, London	HZM.15.4400	Skin	1
P. phoenicurus	Male	23.05.1930	Bern, Switzerland	HZM, London	136.HZM.64386	Skin	1
P. phoenicurus	Male	22.05.1936	Bern, Switzerland	HZM, London	HZM.8.4393	Skin	1
P. phoenicurus	Male	08.08.1970	Sweden	HZM, London	HZM.64.4449	Skin	2
P. phoenicurus	Male	10.07.1933	Bern, Switzerland	HZM, London	HZM.16.4401	Skin	1
P. phoenicurus	Male	24.06.1923	Altai, West-Siberia, Russia	NHM, Copenhagen	31.918	Skin	1
P. phoenicurus	Male	14.05.1914	Tomsk, West-Siberia, Russia	NHM, Copenhagen	31.906	Skin	1
P. phoenicurus	Male	29.05.1896	Tomsk, West-Siberia, Russia	NHM, Copenhagen	31.903	Skin	1
P. phoenicurus	Male	04.06.1923	Altai, West-Siberia, Russia	NHM, Copenhagen	31.917	Skin	1

P. phoenicurus	Male	28.05.1923	Altai, West-Siberia, Russia	NHM, Copenhagen	31.916	Skin	1
P. phoenicurus	Male	02.07.1914	Tomsk, West-Siberia, Russia	NHM, Copenhagen	31.908	Skin	1
P. phoenicurus	Male	30.08.1917	Kainsk, West-Siberia, Russia	NHM, Copenhagen	31.912	Skin	1
P. phoenicurus	Male	06.09.1917	Kainsk, West-Siberia, Russia	NHM, Copenhagen	31.913	Skin	1
P. phoenicurus	Male	August. 1938	North of Tunis, Tunisia	NHM, Copenhagen	28.807	Skin	1
P. phoenicurus	Male	27.11.1910	Sørvaag, Faroe Islands	NHM, Copenhagen	71.590	Skin	1
P. phoenicurus	Male	01.05.1898	Sørvaag, Faroe Islands	NHM, Copenhagen	71.588	Skin	1
P. phoenicurus	Male	01.05.1898	Sørvaag, Faroe Islands	NHM, Copenhagen	71.587	Skin	1
P. phoenicurus	Male	Breeding plumage	Dalmatien prap. Pregl. Croatia	NHM, Copenhagen	31.902	Skin	1
P. phoenicurus	Male	22.05.1956	Cambridge-Shire, England	NHM, Tring, London	1879.4.5.153	Skin	1
P. phoenicurus	Male	05.08.1935	Korab Mountains, Macedonia	NHM, Tring, London	1936:12:15:37.	Skin	2
P. phoenicurus	Male	02.08.1893	Telschen, Germany	NHM, Tring, London	1934.1.1.4398	Skin	1
P. phoenicurus	Male	Breeding plumage	Nalchik Trek District, Caucasus,Russia	NHM, Tring, London	1902.12.7.32	Skin	2
P. phoenicurus	Male	Breeding plumage	Nalchik Trek District, Caucasus,Russia	NHM, Tring, London	1902.12.7.33	Skin	2

P. phoenicurus	Male	21.05.1876	Kavak, Asia Minor, Turkey	NHM, Tring, London	1898.9.1.2034	Skin	1
P. phoenicurus	Female	10.07.1919	Azrou, Lesser Atlas Central, Morocco	NHM, Tring, London	1919.12.11.209	Skin	1
P. phoenicurus	Female	23.05.1933	Kent, England	HZM, London	HZM.97.4481	Skin	2
P. phoenicurus	Male	10.08.1915	Fasnakyle, Scotland	HZM, London	HZM.62.4447/202	Skin	1
P.p.samamisicus	Unknown	22.07.1940	Mazandaran, Iran	Field Museum, Chicago	238886	Skin	1
P.p.samamisicus	Juvenil	16.08.1940	Mazandaran, Iran	Field Museum, Chicago	238888	Skin	1
P.p.samamisicus	Juvenil	17.08.1940	Bardu Forest, Khoresan, Iran	Field Museum, Chicago	238890	Skin	1
P.p.samamisicus	Juvenil	07.07.1940	Durud, Luristan, Iran	Field Museum, Chicago	238898	Skin	2
P.p.samamisicus	Unknown	27.05.1940	Durud, Luristan, Iran	Field Museum, Chicago	238912	Skin	2
P.p.samamisicus	Female	21.05.1876	Karaku, Pakistan	NHM, Tring, London	BMNH 1898.9.1.2033	Skin	1
P. alaschanicus	Male	May, 1895	Southern Kuku-Noor range, China	NHM, Tring, London	BMNH 1965.M.11028	Skin	
P. alaschanicus	Male	April, pre 1898	Chuan Che-Sup, China	NHM, Tring, London	BMNH 1898.9.1.2203	Skin	
P. hodgsoni	Male	05.05.1936	Chayul Valley, Tibet	NHM, Tring	BMNH 1937.1.17.463	Skin	

P. hodgsoni	Male	03.04.1947	Kongpo, Tibet	NHM, Tring, London	BMNH 1948.27.199	Skin
P.frontalis	Female	23.05.1925	Kashmir, India	Field museum, Chicago	60534	Skin
P. frontalis	Male	23.05.1925	Kashmir, India	Field museum, Chicago	60535	Skin
P. caeruleocephalus	Male	17.07.1948	Uttar Predesh Kumaum, India	Field museum, Chicago	238999	Skin
P. caeruleocephalus	Male	20.07.1948	Uttar Predesh Kumaum, India	Field museum, Chicago	239000	Skin
P. schisticeps	Juvenile	Unknown	Szechwan, China	Field museum, Chicago	68768	Skin
P. schisticeps	Female	17.04.1929	Szechwan, China	Field museum, Chicago	68769	Skin
P.moussieri	Male	16.05.1904	Near Blidah, Algeria	NHM, Tring, London	BMNH 1934.1.1.4449	Skin
P.moussieri	Male	18.05.1897	Amsmiz, South Morocco	NHM, Tring, London	BMNH 1949.Whi.1.3314	Skin

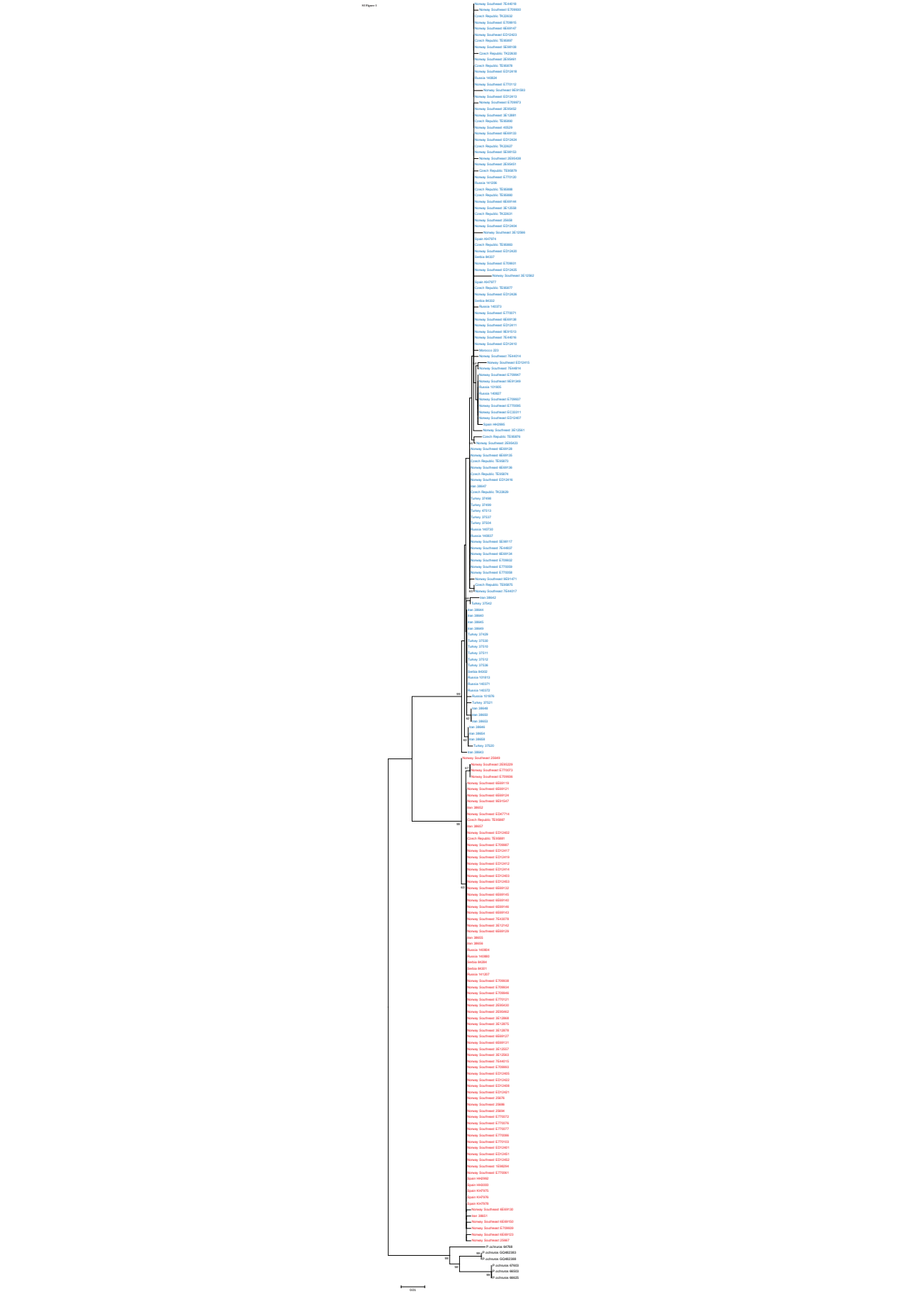
NHM, Oslo = Natural History Museum of Oslo, Norway; NHM Copenhagen = Natural History Museum of Copenhagen, Denmark; HZM, London = Harrison Zoological Museum, London, UK; NHM, Tring, London = Natural History Museum of London, Tring, UK; Field museum, Chicago= Field museum, Chicago, USA

SI Table 3. Details of the sequences downloaded from Genbank.

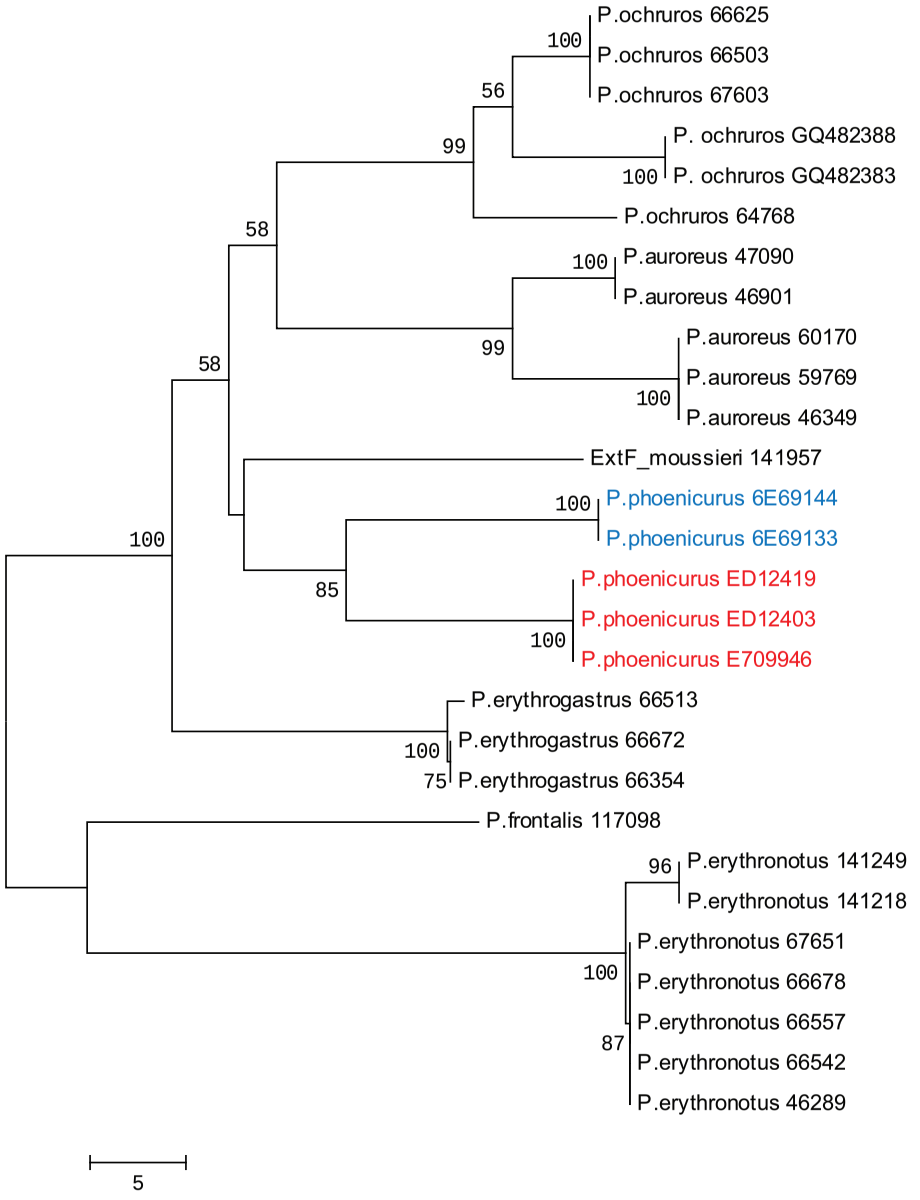
Species	Sex	Date captured	Place captured	Collection/ sampled by	Journal nr	Accession nr
P. ochruros	Male	07.09.1999	Krasnoyarsk Territory, Russia	Burke Museum, Washington	64768	GQ482386
P. ochruros	Male	06.08.1999	Tuva, Russia	Burke Museum, Washington	67603	GQ482385
P. ochruros	Male	20.06.2000	Tuva, Russia	Burke Museum, Washington	66503	GQ482384
P. ochruros	Male	11.06.2000	Tuva, Russia	Burke Museum, Washington	66625	GQ482387
P. ochruros	Female		Kaliningradskaya Oblast, Russia	ZMMU		GQ482383
P. ochruros	Male		Kaliningradskaya Oblast	ZMMU		GQ482388
P. auroreus	Male	12.05.1998	Dornod Aymag Mongolia	Burke Museum, Washington	59769	GQ482372
P. auroreus	Male	17.06.1993	Buryatiya Respublika, Russia	Burke Museum, Washington	46349	GQ482374
P. auroreus	Male	16.06.1993	Khabarovsk Territory, Russia	Burke Museum, Washington	47090	GQ482373
P. auroreus	Male	06.08.1998	Tov, Mongolia	Burke Museum, Washington	60170	GQ482370
P. auroreus	Male	21.06.1993	Khabarovsk Territory, Russia	Burke Museum, Washington	46901	GQ482371
P. erythrogastrus	Male	22.06.2000	Tuva, Russia	Burke Museum, Washington	66354	GQ482377
P. erythrogastrus	Male	23.06.2000	Tuva, Russia	Burke Museum, Washington	66513	GQ482375

P. erythrogastrus	Male	22.06.2000	Tuva, Russia	Burke Museum, Washington	66672	GQ482376
P. erythronotus	Male	06.05.1993	Republic of Gorno-Altay, Russia	Burke Museum, Washington	46289	GQ482382
P. erythronotus	Male	29.06.1999	Tuva, Russia	Burke Museum, Washington	67651	GQ482380
P. erythronotus	Male	29.06.1999	Tuva, Russia	Burke Museum, Washington	66542	GQ482378
P. erythronotus	Male	07.02.2000	Tuva, Russia	Burke Museum, Washington	66557	GQ482381
P. erythronotus	Male	24.06.2000	Tuva, Russia	Burke Museum, Washington	66678	GQ482379

ZMMU = Zoological Museum of Moscow University, GQ numbers are Genbank accession numbers


S1 Figure 1
Norway Southeast 7644016
Norway Southeast 6709693
Czech Republic TK22632
Norway Southeast 6709615
Norway Southeast 6E69147
Norway Southeast ED12623
Czech Republic TE66697
Norway Southeast 3E69138
Czech Republic TK22630
Norway Southeast 3E66661
Czech Republic TE66678
Norway Southeast ED12618
Russia 143624
Norway Southeast 6770112
Norway Southeast 6E70583
Norway Southeast ED12613
Norway Southeast 6709973
Norway Southeast 3E64652
Norway Southeast 3E13681
Czech Republic TE66640
Norway Southeast 40529
Norway Southeast 6E69133
Norway Southeast ED12634
Czech Republic TK22627
Norway Southeast 6E68153
Norway Southeast 26E6458
Norway Southeast 3E66451
Czech Republic TE66679
Norway Southeast 6770120
Russia 141238
Czech Republic TE66688
Czech Republic TE66680
Norway Southeast 6E69144
Norway Southeast 3E13638
Czech Republic TK22631
Norway Southeast 26658
Norway Southeast ED12634
Norway Southeast 3E12566
Spain KH7974
Czech Republic TE66683
Norway Southeast ED12630
Serbia 84337
Norway Southeast 6709631
Norway Southeast ED12635
Norway Southeast 3E12562
Spain KH7977
Czech Republic TE66677
Norway Southeast ED12626
Serbia 84332
Russia 140372
Norway Southeast 6770571
Norway Southeast 6E69136
Norway Southeast ED12611
Norway Southeast 6E67013
Norway Southeast 7644016
Norway Southeast 6E70110
Morocco 223
Norway Southeast 7644014
Norway Southeast ED12615
Norway Southeast 7644014
Norway Southeast 6709567
Norway Southeast 6E67368
Russia 101905
Russia 140927
Norway Southeast 6709697
Norway Southeast 6770095
Norway Southeast 6CD3311
Norway Southeast 6D12467
Spain H42065
Norway Southeast 3E12561
Czech Republic TE66676
Norway Southeast 26365422
Norway Southeast 6E69128
Norway Southeast 6E69135
Czech Republic TE66672
Norway Southeast 6E69136
Czech Republic TE66674
Norway Southeast ED12616
Iran 38647
Czech Republic TK22629
Turkey 37498
Turkey 37498
Turkey 47513
Turkey 37537
Turkey 37534
Russia 140732
Russia 140927
Norway Southeast 6E69117
Norway Southeast 7644637
Norway Southeast 6E69134
Norway Southeast 6709692
Norway Southeast 6770059
Norway Southeast 6770058
Norway Southeast 66914171
Czech Republic TE66675
Norway Southeast 7644017
Iran 38642
Turkey 37542
Iran 38644
Iran 38643
Iran 38645
Iran 38649
Turkey 37439
Turkey 37530
Turkey 37510
Turkey 37511
Turkey 37512
Turkey 37536
Serbia 84202
Russia 101913
Russia 140371
Russia 140372
Russia 101876
Turkey 37521
Iran 38648
Iran 38650
Iran 38653
Iran 38646
Iran 38654
Iran 38658
Turkey 37520
Iran 38642
Norway Southeast 25669
Norway Southeast 3E66229
Norway Southeast 6770073
Norway Southeast 6709696
Norway Southeast 66695119
Norway Southeast 66695121
Norway Southeast 66695124
Norway Southeast 6691647
Iran 38652
Norway Southeast ED47714
Czech Republic TE66687
Iran 38657
Norway Southeast ED12452
Czech Republic TE66691
Norway Southeast 6709687
Norway Southeast ED12417
Norway Southeast ED12419
Norway Southeast ED12412
Norway Southeast ED12414
Norway Southeast ED12453
Norway Southeast ED12463
Norway Southeast 66695132
Norway Southeast 66695145
Norway Southeast 66695140
Norway Southeast 66695146
Norway Southeast 66695143
Norway Southeast 7643078
Norway Southeast 3E13742
Norway Southeast 66695129
Iran 38655
Iran 38656
Russia 140904
Russia 140880
Serbia 84284
Serbia 84301
Russia 141207
Norway Southeast 6709698
Norway Southeast 6709634
Norway Southeast 6709646
Norway Southeast 6770121
Norway Southeast 26365435
Norway Southeast 26365462
Norway Southeast 3E12868
Norway Southeast 3E12875
Norway Southeast 3E12878
Norway Southeast 66695127
Norway Southeast 66695131
Norway Southeast 3E12557
Norway Southeast 3E12563
Norway Southeast 7644015
Norway Southeast 6709693
Norway Southeast ED12455
Norway Southeast ED12422
Norway Southeast ED12458
Norway Southeast ED12421
Norway Southeast 23676
Norway Southeast 25686
Norway Southeast 25694
Norway Southeast 6770072
Norway Southeast 6770076
Norway Southeast 6770077
Norway Southeast 6770088
Norway Southeast 6770103
Norway Southeast ED12401
Norway Southeast ED12451
Norway Southeast ED12452
Norway Southeast 19365244
Norway Southeast 6770061
Spain H42062
Spain H42020
Spain KH7975
Spain KH7976
Spain KH7978
Norway Southeast 66695130
Iran 38651
Norway Southeast 66695150
Norway Southeast 67099609
Norway Southeast 66695133
Norway Southeast 25687
P. ochracea 64768
P. ochracea GO462393
P. ochracea GO462388
P. ochracea 67629
P. ochracea 66503
P. ochracea 66626
0.01

SI Figure 2


P.ochruros 66625
P.ochruros 66503
P.ochruros 67603
P. ochruros GQ482388
P. ochruros GQ482383
P.ochruros 64768
P.auroreus 47090
P.auroreus 46901
P.auroreus 60170
P.auroreus 59769
P.auroreus 46349
ExtF_moussieri 141957
P.phoenicurus 6E69144
P.phoenicurus 6E69133
P.phoenicurus ED12419
P.phoenicurus ED12403
P.phoenicurus E709946
P.erythrogastrus 66513
P.erythrogastrus 66672
P.erythrogastrus 66354
P.frontalis 117098
P.erythronotus 141249
P.erythronotus 141218
P.erythronotus 67651
P.erythronotus 66678
P.erythronotus 66557
P.erythronotus 66542
P.erythronotus 46289
100
56
99
100
58
100
99
58
100
100
85
100
100
75
96
100
87
5

II

Ecology and Evolution

Open Access

Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*)

Kjersti S. Kvie[1,2], Silje Hogner[1], Leif Aarvik[1], Jan T. Lifjeld[1] & Arild Johnsen[1]

[1]Natural History Museum, University of Oslo, P.O. Box 1172, Blindern, N-0318, Oslo, Norway
[2]Department of Basic Sciences and Aquatic Medicine, Norwegian School of Veterinary Science, P.O. Box 8146 Dep, N-0033, Oslo, Norway

Keywords
DNA barcoding, *Epirrita autumnata*, mtDNA, nDNA, selective sweep, *Wolbachia*.

Correspondence
Kjersti S. Kvie, Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, N-0318 Oslo, Norway. Tel: +47 22964793; Fax: +47-22851837; E-mail: kjersti.kvie@nvh.no

Funding Information
This study was funded by Natural History Museum, University of Oslo.

Received: 10 October 2012; Revised: 29 October 2012; Accepted: 1 November 2012

Ecology and Evolution **2013; 3(1): 126–144**

doi: 10.1002/ece3.434

Abstract

Deep sympatric intraspecific divergence in mtDNA may reflect cryptic species or formerly distinct lineages in the process of remerging. Preliminary results from DNA barcoding of Scandinavian butterflies and moths showed high intraspecific sequence variation in the autumnal moth, *Epirrita autumnata*. In this study, specimens from different localities in Norway and some samples from Finland and Scotland, with two congeneric species as outgroups, were sequenced with mitochondrial and nuclear markers to resolve the discrepancy found between mtDNA divergence and present species-level taxonomy. We found five COI sub-clades within the *E. autumnata* complex, most of which were sympatric and with little geographic structure. Nuclear markers (ITS2 and Wingless) showed little variation and gave no indications that *E. autumnata* comprises more than one species. The samples were screened with primers for *Wolbachia* outer surface gene (*wsp*) and 12% of the samples tested positive. Two *Wolbachia* strains were associated with different mtDNA sub-clades within *E. autumnata*, which may indicate indirect selection/selective sweeps on haplotypes. Our results demonstrate that deep mtDNA divergences are not synonymous with cryptic speciation and this has important implications for the use of mtDNA in species delimitation, like in DNA barcoding.

Introduction

Species are often regarded as basic units of evolution and correct species delimitation serves as a backbone in most biological studies (Mayr 1982; Roe and Sperling 2007). However, the number of described species is a small portion of the estimated extant number and there is a need for an increased ability to identify and discriminate species (Blaxter 2004; Silva-Brandao et al. 2009). For the last three decades, mitochondrial DNA has been extensively used (Ballard and Whitlock 2004) and proven to be an important tool in species delimitation as it possesses biological properties making it suitable as a marker for molecular biodiversity (Moore 1995; Hebert et al. 2003).

A universal system for rapid, inexpensive species identification applicable for any life stage, DNA barcoding, has been proposed by Hebert et al. (2003). The ambition behind DNA barcoding is identification by sequencing of short standardized gene regions in order to assign unknown individuals to species and to enhance the discovery of new species. The assumptions underlying DNA barcoding are that every species have sets of unique barcode sequences and hence constitutes monophyletic clades and that genetic variation between species exceeds the variation within species (Hebert et al. 2003). Nevertheless, there are examples of deep intraspecific divergences in mtDNA, also in sympatric populations of animal groups such as birds (Omland et al. 2000; Johnsen et al. 2010; Hogner et al. in press), beetles (Schulenburg et al. 2002; Avtzis et al. 2008), and spiders (Chang et al. 2007). There are several possible explanations for high intraspecific variation. First, this pattern may reflect the presence of cryptic species. The exploration of cryptic species within the Skipper butterfly, *Astraptes fulgerator*, performed by Hebert et al. (2004) is a well-known example. By combining DNA barcoding with information about ecology and morphology of *A. fulgerator*, at least 10, largely sympatric cryptic species were revealed (but see Brower 2006). Second, demographic effects like isolation will cause differentiation between isolated populations by the accumulation of mutations over time. The differentiation may

then reflect early stages of speciation. Secondary admixture of allopatrically evolved populations will in many cases result in gene trees with pronounced phylogenetic gaps between branches (Avise 2000). However, haplotype loss due to genetic drift (i.e., lineage sorting) will over time make a population monophyletic for a single gene lineage (Beebee and Rowe 2004). As lineage sorting is more prominent in small populations, the number of haplotypes maintained in a population is a function of current and historic effective population sizes. In closely related species, allele fixation often fails to complete and they will in these cases share ancestral polymorphisms resulting in discordance between gene trees and species trees (Moore 1995; Beebee and Rowe 2004). For nuclear DNA, when reproductive barriers do not evolve in allopatry and if secondary contact is obtained, variation must be maintained by factors opposing gene flow (e.g., geography and ecology). This is because gene flow will homogenize the nuclear genome over time (Futuyma 2005). Third, introgression by hybridization between closely related species can cause mtDNA to show a different gene genealogy than most other genes in the species in question. As the gene genealogy resulting from introgression is very similar to that expected by ancestral polymorphism and incomplete linage sorting (Ballard and Whitlock 2004), distinguishing between isolation and ancient hybridization can be very difficult. Finally, interpretation of mitochondrial genetic diversity may be hampered by the presence of heritable endoparasitic symbionts and in some cases result in incongruence between nDNA and mtDNA (Linares et al. 2009). Among the most widespread are bacteria from the genus *Wolbachia* (Alphaproteobacteria: Rickettsiales) (Russell et al. 2009). It has long been recognized that endoparasitic symbionts are prevalent among arthropods and that these organisms may have an important role in arthropod evolution as they can cause a number of reproductive alterations in their host, the most common being cytoplasmic incompatibility (Rousset et al. 1992; Werren 1997; Hurst et al. 1999; Hurst and Jiggins 2005; Narita et al. 2009). Male-killing parthenogenesis and feminization of genetic males are other alterations documented in arthropods (Rousset et al. 1992; Grandjean et al. 1993; Werren et al. 1995; Werren 1997; Jiggins 2003; Hurst and Jiggins 2005). The effects of inherited symbionts can be mistaken as evidence for population structure and admixture, as an mtDNA genealogy with deep internal branches could be the result of multiple selective sweeps from different *Wolbachia* strains, rather than a population being large and old or because of secondary admixture (Hurst and Jiggins 2005). Nevertheless, analysis and comparison of sequence data from both mtDNA and nDNA should help distinguishing between demographic effects and indirect selection on mtDNA by parasitic bacteria in an infected population (Rokas et al. 2001; Raychoudhury et al. 2010).

The genus *Epirrita* constitutes nine species (Scoble 1999), of which three are distributed in Norway (Aarvik et al. 2009). These are the autumnal moth, *Epirrita autumnata,* pale November moth, *Epirrita christyi,* and November moth, *Epirrita dilutata*. *E. autumnata* (Fig. 1) is distributed from Japan and Manchuria through Mongolia, Siberia, and Caucasus, to Western Europe and from the northern parts of Scandinavia to the Mediterranean (Skou 1984). The subspecies *E. autumnata omissa* and *E. autumnata henshawi* are found in North America (Tenow 1972; Scoble 1999). The larvae feed on deciduous trees, especially on birch (*Betula*), alder (*Alnus*), and willow (*Salix)* and have cyclic outbursts with ~9- to10-year intervals (Tenow 1972; Aarvik et al. 2009). In periods with high larvae densities, it can defoliate and seriously harm the mountain birch (*Betula pubescens* ssp. *czerepanovii*) forests (Ruohomaki et al. 2000; Jepsen et al. 2008; Yang et al. 2008). As a consequence of the moths' cyclical population dynamics, northern populations of *E. autumnata* may experience present-day bottlenecks as outbreaks are followed by collapse in population size and subsequent decline in genetic variability. Hence, one might expect to find relatively low levels of genetic variation within this species (Futuyma 1998; Snäll et al. 2004). However, preliminary results from DNA barcoding of Scandinavian moths and butterflies (Lepidoptera) revealed discrepancy between present division to species and sequence divergence in the genus *Epirrita* (Johnsen, Aarvik & Lifjeld, unpublished data). In particular, high sequence variation clustered in several well-defined haplogroups within sympatric *E. autumnata* suggested that this might be a complex of cryptic species.

The main aims of this study were to examine the relatively high mtDNA variation found within *E. autumnata*,

Figure 1. The study species, *Epirrita autumnata*. Photo: Svein Bekkum.

describe the degree of sympatry among haplogroups within Norwegian populations of this species and compare the variation at mitochondrial (Cytochrome *c* Oxidase subunit 1, CO1) and nuclear (Internal Transcribed Spacer 2, ITS2 and Wingless) loci. In particular, we wanted to investigate four possible explanations for high intraspecific mtDNA variation: (1) presence of cryptic species; (2) historic isolation and secondary contact; (3) introgression from a related species; and (4) *Wolbachia* infections associated with different haplogroups. First, if the high mtDNA diversity reflects cryptic species, we predict congruence between divergence in mtDNA and nDNA sequence data, given that there has been sufficient time for divergence. Second, if the pattern is due to isolation and secondary contact, we predict higher differentiation in mtDNA compared with nDNA because the former has a relatively high evolutionary rate (5–10 times higher than single copy nDNA) (Avise 1986). Furthermore, depending on the amount of time since range expansions and secondary contact, we expect some degree of mtDNA- and nDNA structure, reflecting the demographic history and original geographic distribution of the lineages, again with higher degree of structure in mtDNA. Third, if ancient introgression by hybridization caused the differentiation in *E. autumnata* mtDNA, the same predictions as for historic isolation with secondary contact will apply. However, if introgression occurred recently, we would expect to find overlapping haplotypes with closely related species (e.g., *E. dilutata* and/or *E. christyi*). Finally, if *Wolbachia* infections have affected the mtDNA variation within this species, we predict an association between infection status and haplogroups and incongruence between mtDNA and nDNA. The samples were screened for *Wolbachia* infections to evaluate whether *Wolbachia* might have influenced patterns of mitochondrial diversity in *E. autumnata*.

Material and methods

Material examined

A total of 87 moths from the genus *Epirrita* were examined in the course of this study, of which 79 were collected in Norway, five in Finland, and three in Scotland (Appendix, Table A1). The Norwegian moths were sampled from different parts of Norway in the period 1999–2009. The middle leg of each moth was collected and stored in ethanol for DNA extraction and the abdomen was removed from some of the specimens for the purpose of making genital preparations. The rest of the animal was prepared dry and pinned as voucher. Information about the samples is available at the Barcode of Life Data Systems website (http://www.boldsystems.org) in the "NorBOL – Lepidoptera – Epirrita" project. In addition, two Wingless and three *wsp* sequences (see below) were downloaded from GenBank and included in the analysis. Sequences downloaded from GenBank are identified by accession numbers in the phylogenetic trees.

Genetic analysis

DNA extraction

Legs were dried at 50°C and transferred to eppendorf tubes. To speed up lysis, the legs were crushed into smaller pieces. DNA extraction was carried out using the E.Z.N.A tissue kit (Omega Bio-tek Inc, Norcross, GA), according to the manufacturer's protocol. The lysis reaction proceeded overnight and the DNA elution was performed with 100-μL elution buffer.

Amplification

Amplification of a 658 base pair long COI fragment from the COI 5′ region was performed using the primers Lep-F1 (5′-ATTCAACCAATCATAAAGATAT-3′; Hebert et al. 2004) and Lep-R1, (5′-TAAACTTCTGGATGTCCAAAAA-3′ Hebert et al. 2004). In cases where these primers failed to amplify, a second reverse primer EnhLep-R1 (5′-CTCCWCCAGCAGGATCAAAA-3′; Hajibabaei et al. 2006) was used in combination with Lep-F1, targeting a 612-bp fragment of the COI region. The PCR profile used for this marker was as follows: 94°C for 1 min, 94°C for 30 sec, 46°C for 40 sec, 72°C for 1 min, (step 2–4 cycled 5 times), 94°C for 30 sec, 51°C for 40 sec, 72°C for 1 min, (step 5–7 cycled 35 times), and 72°C for 10 min.

A 500- to 514-bp long fragment, depending on the species, from the ITS2 region was amplified using the forward primer ITS3b (5′-GGGTCGATGAAGAACGCAST-3′; Roe and Sperling 2007) and reverse primer ITS4 (5′-TCCTCCGCTTATTGATATGC-3′; White et al. 1990). If these primers failed to amplify, another forward primer, FFA (5′-TGTGAACTGCAGGACACA-3′, Brown et al. 2000) was used. PCR profiles for the ITS2 markers were as follows: 94°C for 2 min, 94°C for 30 sec, 55°C for 1 min, 72°C for 2 min, (step 2-4 cycled 34 times), and 72°C for 10 min.

To amplify a 408-bp long fragment from the Wingless region, a single primer pair was used; LepWG1_f (5′-GARTGYAARTGYCAYGGYATGTCTGG-3′; Brower and DeSalle 1998) and LepWG2_r (5′-ACTICGCRCACCA RTGGAATGTRCA-3′; Brower and DeSalle 1998). The PCR profile was as follows: 95°C for 5 min, 95°C for 1 min, 50°C for 1 min, 72°C for 2 min, (step 2–4 cycled 35 times) and 72°C for 10 min. For some of the Wingless samples, more than one fragment was amplified. In these cases, a

second electrophoresis was performed using 7-μL PCR product and 4-μL loading dye. The PCR product was cut out with scalpel under UV light, cleaned up, and DNA was extracted following the protocol NucleoSpin® Extract II, PCR clean-up/extraction kit (Macherey-Nagel, Düren, Germany). To solubilize the gel slices, 200 μL NT buffer pr. 100 μg gel/PCR product was used.

General *wsp* primers were used to amplify 555–560 bp, depending on the strain, from the *Wolbachia* outer surface gene; *wsp* 81F (5′-TGG TCC AAT AAG TGA TGA AGA-AAC-3′; Braig et al. 1998) and *wsp* 691R (5′-AAA AAT TAA ACG CTA CTC CA-3′; Braig et al. 1998). The following PCR profile was used with the *wsp* primers: 94°C for 1 min, 94°C for 30 sec, 55°C for 40 sec, 72°C for 1 min, (step 2–4 cycled 35 times) and 72°C for 10 min. All *wsp* sequences were cloned in case of multiple infections (see below).

PCR reactions were performed in 10- or 12.5-μL reaction volume. The final concentration of the various chemicals was as follows: 1× buffer, 1.5mM $MgCl_2$, 0.8mM dNTPs, 0.5 mM of the forward and reverse primers, 3% DMSO, 1U/μL Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA), and dH_2O to make up the remaining reaction volume. The DNA template had a final concentration of 15–50 ng. The $MgCl_2$ concentration and/or polymerase concentration was increased when no bands were visible with agarose gel screening. All samples were screened on 1% agarose gel, stained with ethidiumbromide or SYBR- Safe (Invitrogen). *Wolbachia* screenings were performed with positive control. In cases where no bands were visible, a second amplification/screening was performed to confirm the result.

Cloning of *wsp* sequences

Cloning was performed following the TOPO10 Cloning protocol (Invitrogen). The PCR product was heated to 68°C for 10 min before the TOPO cloning reaction was set up. We used 2-μL PCR product and let the reaction incubate for 15 min at room temperature. We used *E.coli* DH5α cells for transformation and the transformed cells were transferred to growth medium (LB agar) containing Kanamycin (100 μg/mL) as selection marker. DNA from 6–8 clones from each individual were picked out and diluted in 6-μL dH_2O (for two individuals only three colonies were obtained, however, they all gave the same result). The samples were then amplified and sequenced as described below using standard M13 primers.

Sequencing

The samples were cleaned for unconsumed primers and nucleotides using Exo-Sap-IT (United States Biochemical, Cleveland, OH), diluted 10 times and incubated at 37°C for 45 min for degradation of excess primers and nucleotides and inactivated at 80°C for 15 min. Cycle sequencing was performed in 10-μl reaction volume, using BigDye v3.1 cycle sequencing kit with 5× BigDye Terminator sequencing buffer (Applied Biosystems, Foster City, CA) and a program following manufacturer's recommendations. Purification was performed using ethanol/EDTA/sodium acetate precipitation. Electrophoresis and data analysis of samples were performed with an ABI 3130×l capillary electrophoresis instrument.

The four regions were sequenced in both directions and the resulting consensus sequences were aligned by ClustalW and manually edited in MEGA version 4.0.2 (Tamura et al. 2007). Using highly conserved primers, there is a risk of co-amplifying non-functional copies of mtDNA (numts) in addition to the targeted mtDNA and numts have been shown to be a source of error by overestimation of unique species inferred from the analysis (Gellissen et al. 1983; Lopez et al. 1994; Song et al. 2008). Careful examination of the sequences can reveal numts based on properties such as indels, frameshift mutations, in-frame stop codons, unexpected nucleotide composition, and systematic double peaks (Song et al. 2008). Alignments generated from the three coding regions (COI, Wingless, and *wsp*) were translated from nucleotide- to amino acid sequences to check for stop codons and frameshift mutations.

Phylogenetic and statistical analyses

A model test was performed on all four data sets using MEGA5 version 5.05 (Tamura et al. 2011) to find the best fit substitution models for the different markers. Neighbor-joining analyses, calculation of genetic distances, and standard errors between the different haplogroups were performed in MEGA5 using the Tamura 3-parameter (Tamura 1992) (COI and *wsp*) – and the Jukes–Cantor algorithm (Jukes and Cantor 1969) (ITS2 and Wingless) with all sites included, the complete deletion option, assuming homogenous pattern among lineages and uniform substitution rates among sites. Bootstrap values were calculated in MEGA5 using 10 000 iterations.

To test for neutrality, DnaSP version 5.10 (Librado and Rozas 2009) was used to compute Tajima's *D* (Tajima 1989). This test is based on the allele frequency spectrum and can be used to infer previous evolutionary and demographic events in the population. Positive values indicate an excess of intermediate-frequency alleles, which might result from balancing selection or bottlenecks, while negative values reflect an excess of rare polymorphisms, which might result from positive selection or a population expansion (Akey et al. 2004). We also calculated the two common measures of nucleotide polymorphism, π, the average number of nucleotide differences per site between

two sequences and θ, the population mutation parameter estimated from the number of segregating sites in the aligned sample of sequences (Nei 1987).

Analysis of molecular variance (AMOVA: Excoffier et al. 1992) and calculation of F_{ST} (Wright 1951) were performed on 53 *E. autumnata* COI sequences of the Norwegian samples using ARLEQUIN version 3.5.1.2 (Excoffier et al. 2005), to investigate how genetic variation was distributed within and between regions. The 14 Norwegian sampling locations were divided into four regions: north ($N = 5$), east ($N = 22$), south ($N = 13$), and west ($N = 13$) (Appendix, Table A1). The analysis was conducted with pairwise difference as distance method.

Results

Mitochondrial and nuclear DNA variation

Translation from nucleotide- to amino acid sequences of the analyzed regions revealed no stop codons, frameshifts, or systematic double peaks and the mtDNA base composition was as expected, with a high A-T content (68%) (Perna and Kocher 1995).

Neighbor-joining analysis of the COI data set showed high intraspecific variation within *E. autumnata*, with 21 haplotypes divided into five distinct haplogroups with varying degree of support at each node, ranging from 63% to 99% (Fig. 2a, only bootstrap values higher than 85% are shown). Standard estimates of nucleotide polymorphism were higher within *E. autumnata* than within *E. christiy* and *E. dilutata* (Table 1). Genetic distance between *E. autumnata* COI haplogroups 1–5 ranged from 1.5% to 4.1% (Table 2). Assuming a COI substitution rate of 1.5–2.3% per million years (Brower 1994; Farrell 2001; Kandul et al. 2004), genetic distance as high as 4.1% (distance between haplogroup 1 and 5) suggests divergence as far back as 1.7–2.7 million years. Interspecific distances among *E. autumnata*, *E. christyi,* and *E. dilutata* ranged from 2.9 to 7.6%. Haplogroup 3 consists of moths from Scotland, while the remaining four groups comprise samples from all four Norwegian regions: north, east, south, and west. This shows a high degree of sympatry of mtDNA linages in the northern *E. autumnata* populations. The Neighbor-joining topology based on the COI data is supported by Minimum evolution and Maximum parsimony analysis generated in MEGA5 (Appendix Fig. A1 and A2). In contrast, the phylogenies based on nuclear loci show far less intraspecific variation. In the phylogenies based on the ITS2- and Wingless data sets, *E. autumnata* constitutes one monophyletic group with 0.5% and no variation, respectively (Fig. 3 and 4). The AMOVA reveals that COI haplotype variation is much higher within regions (98.3%), than between regions (1.7%) (Table 3; overall $F_{ST} = 0.017$, $P = 0.27$). Estimates of Tajima's D were negative, but not significantly different from zero for COI in all three species (Table 1).

Wolbachia screening

Of the 71 samples screened, 17 (12%) tested positive for *Wolbachia*. It was possible to generate *wsp* sequences from 14 of the 17 infected samples and all 14 individuals had only one *wsp* sequence as revealed by cloning. Hence, there was no case of multiple infections. Comparing these sequences with sequences in GenBank matched strains found in various insect taxa, including Lepidoptera (99% match). NJ analysis of the 14 sequences obtained using *wsp* primers combined with sequences downloaded from GenBank, splits the sequences into three clusters with 100% bootstrap support at each node (Fig. 2b). Infections from bacteria in *wsp* group 1 were only found in *E. christyi*, whereas *wsp* group 2 and 3 were found exclusively in *E. autumnata*. Interestingly, the two *wsp* groups infecting *E. autumnata* were associated with different haplogroups: *wsp* group 2 occurred only in *E. autumnata* haplogroup 5, whereas *wsp* group 3 occurred in haplogroups 1 and 4. Given the observed frequency ratio of 1:2 for the two *wsp* groups in *E. autumnata*, the probability that the *Wolbachia* positives within each of three haplogroups ($N = 2$, 3 and 4, respectively) would not show mixed *wsp* genotypes can be estimated to $P = 0.039$. Hence, we conclude that there was a significant association between *Wolbachia* strains and *E. autumnata* haplogroups. Screening for *Wolbachia* also revealed fragments in *E. autumnata* haplogroup 2 (2 individuals). The origin of these bands is unknown as no sequences were obtained from the PCR products. However, it would be interesting to sequence these bands in a future study as they might consist of a more divergent *Wolbachia* strain that requires other suitable primers.

Discussion

We found five distinct mtDNA haplogroups within *E. autumnata* in northern Europe, but little variation was found in the nuclear regions ITS2 and Wingless. High degree of sympatry and little geographic structure in *E. autumnata* haplotype distribution was evident. Twelve percent of the screened samples proved to be infected with *Wolbachia* and there was a close association between particular COI haplogroups within *E. autumnata* and the different *Wolbachia* strains.

Mitochondrial and nuclear incongruence

Combining data sets from unlinked genes should be informative in questions regarding species delineation, as

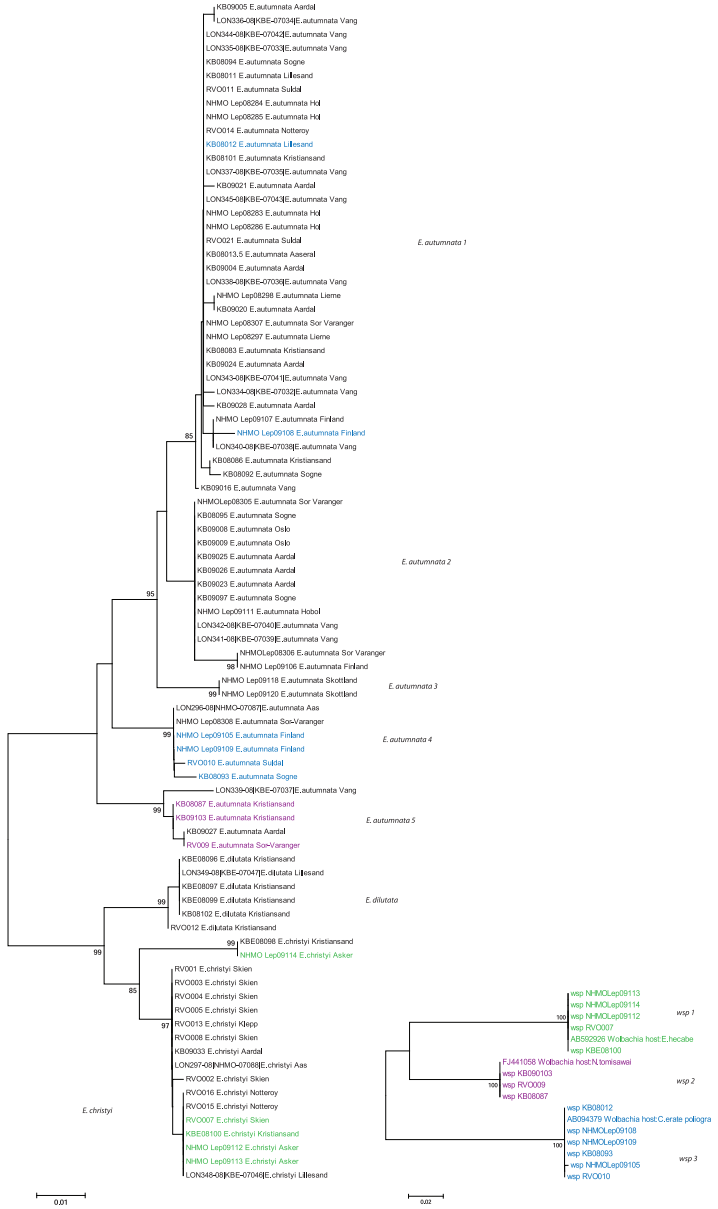


Figure 2. (a) Neighbor-joining analysis of 86 *Epirrita* samples based on the COI gene (Tamura 3-parameter used as substitution model). *E. dilutata* and *E. christyi* are included as out-groups. The five monophyletic groups of distinct *E. autumnata* haplotypes show high intraspecific variation. Bootstrap support (10000 iterations) is shown at each node. (b) Neighbor-joining analysis based on the *wsp* gene (Tamura 3-parameter used as substitution model). Bootstrap support (10000 iterations) is indicated at each node. Infections caused by bacteria represented in *wsp* group 1 are found in *E. christyi*. Infections from *wsp* group 2 are found in *E. autumnata* haplogroup 5, while infections represented in *wsp* group 3 are found in *E. autumnata* group 1 and 4. For each clade, we have added a representative *wsp* sequence (downloaded from Genbank) that has been identified in other Lepidoptera species (*Eurema hecabe*, *Nephopterix tomisawai, Colias erate* subsp. *polygraphus*).

Table 1. Polymorphism statistics for *Epirrita autumnata* (and *Epirrita christyi* and *Epirrita dilutata* for COI) from the COI, Wingless and ITS2 region, and from *Wolbachia* outer surface gene (*wsp*)

	Species	N[1]	L[2]	Tajima's D[3]	π[4]	θ[5]
COI	*E. autumnata*	62	658	-0.77659	0.01674	0.02043
	E. christyi	18	658	-0.89359	0.00693	0.00907
	E. dilutata	6	658	-0.93302	0.00058	0.00076
Wingless	*E. autumnata*	14	409	-1.15524	0.00039	0.00085
ITS2	*E. autumnata*	23	484	1.99999	0.00285	0.542
wsp	*E. autumnata*	17	564	-0.93302	0.10812	0.06573

[1]Number of individuals.
[2]Sequence length.
[3]None of the *D* values significant.
[4]Average pairwise sequence difference per nucleotide (Nei 1987).
[5]Expected heterozygosity per nucleotide (Watterson 1975).

Table 2. Genetic distance (Tamura 3-parameter) between the five *Epirrita autumnata* haplotype groups 1-5 (below diagonal) with standard errors (above diagonal).

	E. a 1	E. a 2	E. a 3	E. a 4	E. a 5
E. autumnata1		0.005	0.007	0.009	0.010
E. autumnata2	0.015		0.007	0.008	0.009
E. autumnata3	0.023	0.022		0.009	0.009
E. autumnata4	0.035	0.033	0.039		0.009
E. autumnata5	0.041	0.036	0.040	0.037	

Table 3. AMOVA of Norwegian *Epirrita autumnata* samples showing haplotype distribution in four selected regions in Norway; north, east, south, and west.

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation
Among regions	3	14.444	0.069 Va	1.710
Within regions	49	194.273	3.965 Vb	98.290
Total	52	208.717	4.034	

unlinked genes are expected to have independent genealogical histories (Maddison 1997). In this study, genetic analysis of gene regions from different genomes (mtDNA and nDNA) gives different estimates of intraspecific variation within *E. autumnata*. The COI region reveals high sympatric intraspecific divergence (Fig. 2a) with genetic distances ranging from 1.5% to 4.1% (Table 2). One might argue that an intraspecific genetic distance of 4% is not high compared with distances found within other taxa. For example, there are several studies on land snails that show a higher degree of intraspecific divergence than that found within *E. autumnata*. However, many of these examples concern isolated and/or morphologically distinct populations (Hayashi and Chiba 2000; Shimizu and Ueshima 2000; Bond et al. 2001; Pinceel et al. 2005). In this context, we want to emphasize that the divergent *E. autumnata* haplogroups occur sympatrically and that the level of genetic distance within *E. autumnata* is comparable to the level of divergence commonly seen between sister species in Lepidoptera (Huemer and Hausmann 2009; Lukhtanov et al. 2009; Hajibabaei 2006; this study). In contrast to the levels of intraspecific variation found in COI, the ITS2 and Wingless regions show little (0.5%) and no variation, respectively (Fig. 3 and 4).

The presence of cryptic species has been suggested to explain high intraspecific divergence in several studies (Hebert et al. 2004; Roe and Sperling 2007; Vaglia et al. 2008). However, the presence of cryptic species predicts divergence in both mtDNA and nDNA and the results from this study show clear incongruence between the two data sets. In addition, preliminary analyses show no obvious intraspecific variation in genital structures and no association between flight period and haplogroups (Kvie & Aarvik, unpublished data). These findings all imply that cryptic speciation is not a likely explanation for high intraspecific mtDNA variation within *E. autumnata*. Nevertheless, using nuclear markers that evolve faster and that are more variable than Wingless and ITS2 might generate a different result than we found in this study. It is a well-known challenge to find nuclear markers that evolve fast enough to separate between cryptic species (Dasmahapatra and Mallet 2006). However, there are examples of studies performed on closely related- and cryptic arthropod species that have used these nuclear markers successfully (Roe and Sperling 2007; Schmitz et al. 2007; Linares et al. 2009; Dincă et al. 2011; Sun et al. 2011).

As the COI data set implies divergence as far back as 1.7–2.7 million years, a possible hypothesis would be separation of *E. autumnata* into different glacial refugia in Pleistocene (2 – 0.01 million years ago). It is a common perception that many extant sister taxa diverged during the cyclic climate in this period (Avise and Walker 1998; Avise 2000; Beebee and Rowe 2004). If the variation found in northern *E. autumnata* mtDNA is a result of separation into several refugia, we would expect some degree of geographic separation restricting gene flow. However, results from the AMOVA (Table 3) shows that most of the genetic variation is found within (98.3%) and not between populations (1.7%). Also, if mtDNA variation in the northern populations is a result of isolation, and we are looking at early stages of speciation, variation should be detectable in both mitochondrial- and nuclear markers (Jiggins and Tinsley 2005). As analysis of *E. autumnata* nDNA only reveals small amounts of variation and the results from the AMOVA show a small degree of variation between the populations, it is not likely that isolation alone can explain the high mtDNA variation found in this study.

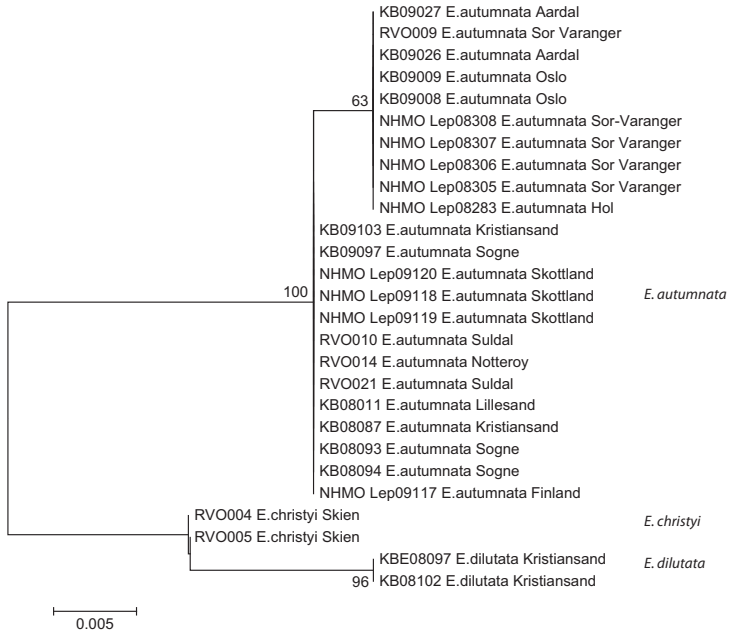


Figure 3. Neighbor-joining analysis of 27 *Epirrita* sequences based on the ITS2 gene (Jukes–Cantor used as substitution model). *E. christyi* and *E. dilutata* are included as out-groups. *E. autumnata* are shown as one monophyletic group with some diversity (< 0.5%). Bootstrap support (10000 iterations) is shown at each node.

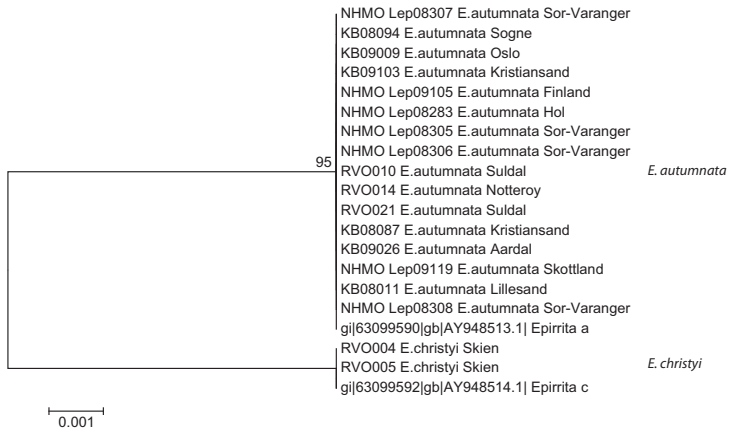


Figure 4. Neighbor-joining analysis of 20 *Epirrita* sequences based on the Wingless gene and with *E. christyi* as out-group (Jukes–Cantor used as substitution model). Bootstrap support (10,000 iterations) is shown at each node.

There are European studies showing opposing results. Snäll et al. (2004) analyzed the mtDNA control region investigating the dispersal of *E. autumnata* females and differentiation between northern- (Norwegian samples) and southern (Finnish samples) *E. autumnata* populations. They found less variation in northern- compared

with southern populations, which they argue might be a result of the northern moths' cyclic population dynamics (Futuyma 1998; Snäll et al. 2004). Their results also revealed moderate levels of divergence between the northern and southern populations. In addition, results from a study performed by Hausmann et al. (2011) showed no variation within Bavarian *E. autumnata* and the published sequences from that project cluster with moths in haplogroup 2 from this study (data not shown). As the two Scottish samples from our study also cluster together in one group (haplogroup 3, Fig. 2a), it is likely that there is a geographic structure at a larger scale and more samples from a wider range should be investigated.

There are no indications of hybridization between the Norwegian *Epirrita* species as no shared haplotypes between *E. christyi*, *E. dilutata,* and *E. autumnata* were found. However, ancient introgression by hybridization or introgression from another extant congeneric cannot be ruled out because this will give similar gene genealogy as ancestral polymorphisms caused by isolation (Ballard and Whitlock 2004). It should be noted that in Lepidoptera, females are the heterogametic sex and that according to Haldane's rule (Haldane 1922), hybrid sterility/inviability will be more severe in the heterogametic sex, thereby reducing the likelihood of heterospecific mtDNA introgression (but see Zakharov et al. 2009). Besides, even if both isolation and ancient introgression by hybridization could explain the origin of intraspecific variation in *E. autumnata*, neither of these processes can explain the high degree of incomplete linage sorting still existing in *E. autumnata* mtDNA. Geography, ecology, and reproductive barriers are all factors that could maintain variation within a species, but we did not find evidence for any of these factors playing a role in this study. Using nuclear markers that evolve faster, performing more thorough morphometric examinations of genitalia, and testing for other ecological and morphological differences like host plant preference and larvae differentiation could give a more solid basis for concluding about these possibilities.

Association between *Wolbachia* infections and COI haplogroups

Screening for *Wolbachia* infections showed infections in *E. autumnata* and in *E. christyi* (Fig. 2a and 2b). This result, combined with the results from NJ analysis of the COI region (Fig. 2a) and the AMOVA (Table 3), resembles those of Schulenburg et al. (2002). They examined Eurasian two-spot ladybirds, *Adalia bipunctata,* infected with endoparasites from the genera *Rickettsia* and *Spiroplasma,* in addition to infections by two distinct strains of *Wolbachia*. Also in this case, did mtDNA sequence analysis show an association between infection status and distribution of haplotypes, but no association between haplotype and geography. However, Shoemaker et al. (2004) showed that *Wolbachia*-infected species tend to have lower levels of mtDNA diversity than uninfected closely related species. Reduced levels of variation are the most commonly documented effect in *Wolbachia*-infected populations (Shoemaker et al. 1999; Dean et al. 2003; Jiggins 2003; Shoemaker et al. 2004). Nevertheless, high levels of diversity in mtDNA may be maintained within a population when infected with bacteria of different strains, as different strains might cause selective sweeps on different haplotypes. The diversity will, in these cases, depend on the number of symbionts the population harbors (Hurst and Jiggins 2005). Symbionts like *Wolbachia* are also known to cause hybrid introgression and possibly balancing selection on cytoplasmic genes and may therefore be an important factor in creating variation within a population or in a species (Jiggins 2003). For example, Jiggins and Tinsley (2005) found significantly elevated levels of mtDNA diversity in infected *Adalia bipunctata* beetles. They argued that the effects of endoparasitic symbionts can be considerably more complex than simple reduction in diversity following a selective sweep. As several samples in this study tested positive for *Wolbachia* and there seems to be an association between haplogroups in *E. autumnata* and infection class, it is possible that the mitochondrial genome of *E. autumnata* has undergone several *Wolbachia* infections and subsequent selective sweeps, maintaining the diversity within this species. However, as the test of selection based on Tajima's *D* gave a non-significant result, we cannot rule out the possibility that drift rather than selective sweeps causes variation to be maintained in this species. Some mtDNA haplotypes and their associated *Wolbachia* variants might be carried to high frequencies because of the cyclical fluctuations in population size in *E. autumnata*.

DNA barcoding Lepidoptera

DNA barcoding has proven to be a useful tool for species identification in a wide range of animal species, including Lepidoptera (Hebert et al. 2004; Hajibabaei 2006; Silva-Brandao et al. 2009; Hausmann et al. 2011, but see Elias et al. 2007; Wiemers and Fiedler 2007). This study shows that sequencing the barcode region is sufficient for discriminating between specimens of Norwegian moths in the genus *Epirrita*, hence fulfilling one of the main objectives of DNA barcoding (species identification of unknown specimens; Hebert et al. 2003). However, our results also demonstrate that delimiting species based on mtDNA divergence alone, whether based on a threshold distance, monophyly, or diagnostic nucleotides (Moritz and Cicero 2004; van Velzen et al. 2012), may lead to erroneous conclusions and inflation of species numbers, supporting

previous critiques of the species discovery aspect of DNA barcoding (Moritz and Cicero 2004; Hickerson et al. 2006). It is becoming increasingly clear that integrating information from several independent genetic loci as well as morphological and/or ecological variation is required for defining new species (DeSalle et al. 2005; Galtier et al. 2009; Damm et al. 2010; Dupuis et al. 2012; Towes and Brelsford 2012). As such, DNA barcoding can be a useful method for initial screening of biodiversity, to discover interesting genetic variation worthy of further study.

Concluding remarks

Analysis of the COI region reveals high divergence within *E. autumnata* compared with the nuclear regions. As 12% of the samples surveyed in this study tested positive for *Wolbachia,* the COI data set should be interpreted with care. Our analyses revealed no association between the distribution of mitochondrial haplotypes and geography. Nevertheless, ecological and morphological factors should be examined more thoroughly to rule out the possibility of the different haplogroups reflecting early stages of speciation. As there seems to be an association between *Wolbachia* infections and mtDNA haplogroups, a likely explanation for the divergences in *E. autumnata* mtDNA is that current populations consist of separate lineages that once evolved in allopatry, without evolving reproductive barriers. At some point, secondary contact is obtained and gene flow reduces variation in the nuclear genome over time, while *Wolbachia* infections contribute to maintain the variation in the mitochondrial genome. The effect of linage sorting also seems prominent as there is one dominant haplogroup (haplogroup 1, Fig. 2a).

From these findings, we conclude that current taxonomy is correct and it is probable that *Wolbachia* contributes to intraspecific mtDNA variation by maintaining less common lineages that normally would have been sorted out.

Acknowledgments

We thank the following persons for contributing with moth samples: Kai Berggren, Geir E. E. Søli, Reidar Voith, Anders Endrestøl, Roy Leverton, and Antti Aalto. We also thank Eirik Rindal for contributing with moth samples and for statistical assistance, Jarl Andreas Anmarkrud for help with data analysis and Torbjørn Ekrem for useful comments. The study was financed by the Natural History Museum of Oslo.

Data accessibility

DNA sequences, GenBank accession: COI: JX260731-JX260816, ITS2: JN225584-JN225610, Wingless: JN225566 -JN225583, Wsp: JX310335-JX310348.

Conflict of Interest

None declared.

References

Aarvik, L., L. O. Hansen, and V. Kononenko. 2009. Norges sommerfugler, Håndbok over Norges dagsommerfugler og nattsvermere. Pp. 174. Norsk entomologisk forening, Naturhistorisk Museum, Oslo (in Norwegian).

Akey, J. M., M. A. Eberle, M. J. Rieder, C. S. Carlson, M. D. Shriver, D. A. Nickerson, et al. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol. 2:1591–1599.

Avise, J. C. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. Phil. Trans. R. Soc. Lond. B 312:325–342.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Pp. 56–59. Harvard University Press, Cambridge, Mass.

Avise, J. C., and D. Walker. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc. R. Soc. Lond. B 265:457–463.

Avtzis, D. N., W. Arthofer, and C. Stauffer. 2008. Sympatric occurrence of diverged mtDNA lineages of *Pityogenes chalcographus* (Coleoptera, Scolytinae) in Europe. Biol. J. Linn. Soc. 94:331–340.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13:729–744.

Beebee, T. J. C., and G. Rowe. 2004. An introduction to molecular ecology. Pp. 167–176. Oxford University Press, Oxford.

Blaxter, M. L. 2004. The promise of a DNA taxonomy. Phil. Trans. R. Soc. B 359:669–679.

Bond, J. E., M. C. Hedin, M. G. Ramirez, and B. D. Opell. 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Mol. Ecol. 10:899–910.

Braig, H. R., W. Zhou, S. L. Dobson, and S. L. O'Neill. 1998. Cloning and characterization of a gene encoding the major surface protein of the bacterial Endosymbiont *Wolbachia pipientis*. J. Bacteriol. 180:2373–2378.

Brower, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl Acad. Sci. USA 91:6491–6495.

Brower, A. V. Z. 2006. Problems with DNA barcodes for species delimitation: 'ten species' of *Astraptes fulgerator* reassessed (Lepidoptera: Hesperiidae). Syst. Biodivers. 4:127–132.

Brower, A. V. Z., and R. DeSalle. 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence

among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. Insect Mol. Biol. 7:73–82.

Brown, B., R. Emberson, and A. Paterson. 2000. Phylogenetic relationships within the genus *Wiseana* (Lepidoptera: Hepialidae). New Zeal J. Zool. 27:1–14.

Chang, J., D. X. Song, and K. Y. Zhou. 2007. Incongruous nuclear and mitochondrial phylogeographic patterns in two sympatric lineages of the wolf spider *Pardosa astrigera* (Araneae: Lycosidae) from China. Mol. Phylogenet. Evol. 42:104–121.

Damm, S., B. Schierwater, and H. Hadrys. 2010. An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology. Mol. Ecol. 19:3881–3893.

Dasmahapatra, K. K., and J. Mallet. 2006. DNA barcodes: recent successes and future prospects. Heredity 97:254–255.

Dean, M. D., K. J. Ballard, A. Glass, and J. W. O. Ballard. 2003. Influence of two *Wolbachia* strains on population structure of East African *Drosophila simulans*. Genetics 165:1959–1969.

DeSalle, R., M. G. Egan, and M. Siddall. 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. Phil. Trans. R. Soc. B Biol. Sci. 360:1905–1916.

Dincă, V., V. A. Lukhtanov, G. Talavera, and R. Vila. 2011. Unexpected layers of cryptic diversity in wood white Leptidea butterflies. Nat. Commun., 2:324.

Dupuis, J. R., A. D. Roe, and F. A. H. Sperling. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol. Ecol. 21:4422–4436.

Elias, M. R., I. Hill, K. R. Willmott, K. K. Dasmahapatra, A. V. Z. Brower, J. Mallet, et al. 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proc. R. Soc. B Biol. Sci. 274:2881–2889.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial - DNA restriction data. Genetics 131:479–491.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. 1:47–50.

Farrell, B. D. 2001. Evolutionary assembly of the milkweed fauna: Cytochrome oxidase I and the age of *Tetraopes* beetles. Mol. Phylogenet. Evol. 18:467–478.

Futuyma, D. J. 1998. Evolutionary biology, 3rd ed. Sinauer Associates, Sunderland, Mass.

Futuyma, D. J. 2005. Evolution. Sinauer Associates, Sunderland, Mass.

Galtier, N., B. Nabholz, S. Glemin, and G. D. D. Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 22:4541–4550.

Gellissen, G., J. Y. Bradfield, B. N. White, and G. R. Wyatt. 1983. Mitochondrial -DNA sequences in the nuclear genome of a locust. Nature 301:631–634.

Grandjean, F., T. Rigaud, R. Raimond, P. Juchault, and C. Soutygrosset. 1993. Mitochondrial-DNA polymorphism and feminizing sex factors dynamics in a natural-population of *Armadillium-vulgare* (Crustacea, Isopoda). Genetica 92:55–60.

Hajibabaei, M., D. H. Janzen, J. M. Burns, W. Hallwachs, and P. D. N. Hebert. 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proc. Natl. Acad. Sci. USA 103:968–971.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet., 12:101–109. Proceedings of the National Academy of Sciences of the United States of America, 103, 968-971.

Hausmann, A., G. Haszprunar, and P. D. N. Hebert. 2011. DNA Barcoding the Geometrid Fauna of Bavaria (Lepidoptera): successes, Surprises, and Questions. PLoS ONE 6:e17134.

Hayashi, M., and S. Chiba. 2000. Intraspecific diversity of mitochondrial DNA in the land snail *Euhadra peliomphala* (Bradybaenidae). Biol. J. Linn. Soc. 70:391–401.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B 270:313–321.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl Acad. Sci. USA 101:14812–14817.

Hickerson, M., C. Meyer, and C. Moritz. 2006. DNA Barcoding Will Often Fail to Discover New Animal Species over Broad Parameter Space. Syst. Biol. 55:729–739.

Hogner, S., T. Laskemoen, J. T. Lifjeld, J. Porkert, O. Kleven, T. Albayrak, et al. (2012) Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart Phoenicurus phoenicurus. Ecol. Evol. (in press). doi:10.1002/ece3.398.

Huemer, P., and A. Hausmann. 2009. A new expanded revision of the European high mountain Sciadia tenebraria. Zootaxa 2117:1–30.

Hurst, G. D. D., and F. M. Jiggins. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc. R. Soc. B 275:1525–1534.

Hurst, G. D. D., F. M. Jiggins, J. H. Graf von der Schulenburg, D. Bertrand, S. A. West, I. I. Goriachea, et al. 1999. Male-killing Wolbachia in two species of insect. Proc. R. Soc. B 266:735–740.

Jepsen, J. U., S. B. Hagen, R. A. Ims, and N. G. Yoccoz. 2008. Climate change and outbreaks of the geometrids *Operophtera brumata* and *Epirrita autumnata* in subarctic birch forest: evidence of a recent outbreak range expansion. J. Anim. Ecol. 77:257–264.

Jiggins, F. M. 2003. Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. Genetics 164:5–12.

Jiggins, F. M., and M. C. Tinsley. 2005. An ancient mitochondrial polymorphism in *Adalia bipunctata* linked to a sex-ratio-distorting bacterium. Genetics 171: 1115–1124.

Johnsen, A., E. Rindal, P. G. P. Ericson, D. Zuccon, K. C. R. Kerr, M. Y. Stoeckle, et al. 2010. DNA barcoding of Scandinavian birds reveals divergent lineages in trans-Atlantic species. J. Ornithol. 151:565–578.

Jukes, T. H., and C. R. Cantor. 1969. *Evolution of protein molecules*. Mammalian Protein Metabolism, Pp. 21–132. Academic Press, New York.

Kandul, N. P., V. A. Lukhtanov, A. V. Dantchenko, J. W. S. Coleman, C. H. Sekercioglu, D. Haig, et al. 2004. Phylogeny of *Agrodiaetus* (Hubner 1822) (Lepidoptera : Lycaenidae) inferred from mtDNA sequences of COI and COII and nuclear sequences of EF1-alpha: Karyotype diversification and species radiation. Syst. Biol. 53:278–298.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions trough comparative studies of nucleotide sequences. J. Mol. Evol. 16:11–120.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Linares, M. C., I. D. Soto-Calderon, D. C. Lees, and N. M. Anthony. 2009. High mitochondrial diversity in geographically widespread butterflies of Madagascar: a test of the DNA barcoding approach. Mol. Phylogenet. Evol. 50:485–495.

Lopez, J. V., N. Yuhki, R. Masuda, W. Modi, and S. J. Obrien. 1994. Numt, a recent transfer and tandem amplification of mitochondrial-DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39:174–190.

Lukhtanov, V. A., A. Sourakov, E. V. Zakharov, and P. D. N. Hebert. 2009. DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. Mol. Ecol. Resour. 9:1302–1310.

Maddison, W. P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

Mayr, E. 1982. The growth of biological thought: diversity, evolution, and inheritance. Belknap Press of Harvard University Press, Cambrigde, Massachusetts.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation- mitochondrial-genetrees versus nuclear-gene trees. Evolution 49:718–726.

Moritz, C., and C. Cicero. 2004. DNA Barcoding: promise and Pitfalls. PLoS Biol. 2:1529–1531.

Narita, S., Y. Shimajiri, and M. Nomura. 2009. Strong cytoplasmic incompatibility and high vertical transmission rate can explain the high frequencies of *Wolbachia* infection in Japanese populations of *Colias erate poliographus* (Lepidoptera: Pieridae). Bull. Entomol. Res. 99:385–391.

Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ, Press, New York.

Omland, K. E., C. L. Tarr, W. I. Boarman, J. M. Marzluff, and R. C. Fleischer. 2000. Cryptiv genetic variation and paraphyly in ravens. Proc. R. Soc. Lond. B 267:2475–2482.

Perna, N. T., and T. D. Kocher. 1995. Patterns of nucleotide composition at a fourfold degenerate sites of animal motochondrial genomes. J. Mol. Evol. 41:353–358.

Pinceel, J. A. N., K. Jordaens, M. Pfenninger, and T. Backeljau. 2005. Rangewide phylogeography of a terrestrial slug in Europe: evidence for Alpine refugia and rapid colonization after the Pleistocene glaciations. Mol. Ecol. 14:1133–1150.

Raychoudhury, R., B. K. Grillenberger, J. Gadau, R. Bijlsma, L. van de Zande, J. H. Werren, et al. 2010. Phylogeography of *Nasonia vitripennis* (Hymenoptera) indicates a mitochondrial-*Wolbachia* sweep in North America. Heredity 104:318–326.

Roe, A. D., and F. A. H. Sperling. 2007. Population structure and species boundary delimitation of cryptic *Dioryctria* moths: an integrative approach. Mol. Ecol. 16:3617–3633.

Rokas, A., R. J. Atkinson, G. S. Brown, S. A. West, and G. N. Stone. 2001. Understanding patterns of genetic diversity in the oak gallwasp *Biorhiza pallida:* demographic history or a *Wolbachia* selective sweep? Heredity 87:294–304.

Rousset, F., D. Bouchon, B. Pintureau, P. Juchault, and M. Solignac. 1992. *Wolbachia* endosymbionts responsible for various alterations of sexuality in arthropods. Proc. R. Soc. Lond. B 250:91–98.

Ruohomaki, K., M. Tanhuanpaa, M. P. Ayres, P. Kaitaniemi, T. Tammaru, and E. Haukioja. 2000. Causes of cyclicity of *Epirrita autumnata* (Lepidoptera, Geometridae): grandiose theory and tedious practice. Popul. Ecol. 42:211–223.

Russell, J. A., B. Goldman-Huertas, C. S. Moreau, L. Baldo, J. K. Stahlhut, J. H. Werren, et al. 2009. Specialization and geographic isolation among *Wolbachia* symbionts from ants and Lycaenid butterflies. Evolution 63:624–640.

Schmitz, P., A. Cibois, and B. Landry. 2007. Molecular phylogeny and dating of an insular endemic moth radiation inferred from mitochondrial and nuclear genes: the genus *Galagete* (Lepidoptera: Autostichidae) of the Galapagos Islands. Mol. Phylogenet. Evol. 45:180–192.

Schulenburg, J., G. D. D. Hurst, D. Tetzlaff, G. E. Booth, I. A. Zakharov, and M. E. N. Majerus. 2002. History of infection with different male-killing bacteria in the two-spot ladybird beetle *Adalia bipunctata* revealed through mitochondrial DNA sequence analysis. Genetics 160:1075–1086.

Scoble, M. J. 1999. Geometrid moths of the world: a catalogue. Pp. 290–291. Natural History Museum, London.

Shimizu, Y., and R. Ueshima. 2000. Historical biogeography and interspecific mtDNA introgression in *Euhadra peliomphala* (the Japanese land snail). Heredity 85:84–96.

Shoemaker, D. D., V. Katju, and J. Jaenike. 1999. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. Evolution 53:1157–1164.

Shoemaker, D. D., K. A. Dyer, M. Ahrens, K. McAbee, and J. Jaenike. 2004. Decreased diversity but increased substitution rate in host mtDNA as a consequence of *Wolbachia* endosymbiont infection. Genetics 168:2049–2058.

Silva-Brandao, K. L., M. L. Lyra, and A. V. L. Freitas. 2009. Barcoding Lepidoptera: current situation and perspectives on the usefulness of a contentious technique. Neotrop. Entomol. 38:441–451.

Skou, P. 1984. Nordens målere: håndbog over de danske fennoskanadiske arter af Drepanidae og Geometridae (Lepidoptera) Pp. 117–119. Fauna Bøger, København.

Snäll, N., K. Huoponen, I. Saloniemi, M. L. Savontaus, and K. Ruohomaki. 2004. Dispersal of females and differentiation between populations of *Epirrita autumnata* (Lepidoptera: Geometridae) inferred from variation in mitochondrial DNA. Eur. J. Entomol. 101:495–502.

Song, H., J. E. Buhay, M. F. Whiting, and K. A. Crandall. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl Acad. Sci. USA 105:13486–13491.

Sun, X. J., J. H. Xiao, J. M. Cook, G. Feng, and D. W. Huang. 2011. Comparisons of host mitochondrial, nuclear and endosymbiont bacterial genes reveal cryptic fig wasp species and the effects of *Wolbachia* on host mtDNA evolution and diversity. BMC Evol. Biol. 11:86.

Tajima, F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tamura, K. 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases. Mol. Biol. Evol. 9:678–687.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596–1599.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731–2739.

Tenow, O. 1972. The outbreaks of Oporinia autumnata Bkh. and Operophthera spp. (Lep., Geometridae) in the Scandinavian mountain chain and northern Finland 1862-1968. Almqvist & Wiksell, Uppsala.

Towes, D. P. L., and A. Brelsford. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21:3907–3930.

Vaglia, T., J. Haxaire, I. J. Kitching, I. Meusnier, and R. Rougerie. 2008. Morphology and DNA barcoding reveal three cryptic species within the *Xylophanes neoptolemus* and *loelia* species-groups (Lepidoptera: Sphingidae). Zootaxa 36:18–36.

van Velzen, R., E. Weitschek, G. Felici, and F. T. Bakker. 2012. DNA barcoding of recently diverged species: relative performance of matching methods. PLoS ONE 7:e30490.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7:256–276.

Werren, J. H. 1997. Biology of *Wolbachia*. Annu. Rev. Entomol. 42:587–609.

Werren, J. H., W. Zhang, and L. R. Guo. 1995. Evolution and phylogeny of *Wolbachia*-reproductive parasites of arthropods. Proc. R. Soc. Lond. B 261:55–63.

White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogeneics. Academic Press Inc., New York.

Wiemers, M., and K. Fiedler. 2007. Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). Front. Zool. 4:8.

Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323–354.

Yang, S. Y., T. Ruuhola, S. Haviola, and M. J. Rantala. 2008. Effects of host-plant shift on immune and other key life-history traits of an eruptive Geometrid, *Epirrita autumnata* (Borkhausen). Ecol. Entomol. 33:510–516.

Zakharov, E. V., N. F. Lobo, C. Nowak, and J. J. Hellmann. 2009. Introgression as a likely cause of mtDNA paraphyly in two allopatric skippers (Lepidoptera: Hesperiidae). Heredity 102(6):590–599.

Appendix: Table A1. Information about the samples surveyed in this study, sampling locations, coordinates and collecting dates. Footnotes behind the municipalities show how the Norwegian *E. autumnata* samples were grouped into four Norwegian regions (used in the AMOVA).

Tissue sample	Species	Country	Municipality	Coordinates	Collecting date
NHMO Lep08305	*Epirrita autumnata*	Norway	Sør-Varanger[1]	69°33′48.53″N, 30°05′11.19″E	10-20.VIII.2008
NHMO Lep08306	*Epirrita autumnata*	Norway	Sør-Varanger[1]	69°27′09.34″N, 30°03′28.97″E	9.IX-2.X.2008
NHMO Lep08307	*Epirrita autumnata*	Norway	Sør-Varanger[1]	69°33′48.53″N, 30°05′11.19″E	10-20.VIII.2008
NHMO Lep08308	*Epirrita autumnata*	Norway	Sør-Varanger[1]	69°33′48.53″N, 30°05′11.19″E	10-20.VIII.2008
RVO009	*Epirrita autumnata*	Norway	Sør-Varanger[1]	69°22′14.45″N,29°40′43.56″E	3.VIII-25.IX.2006
NHMO Lep08297	*Epirrita autumnata*	Norway	Lierne[2]	64°26′44.32″N, 13°42′57.43″E	25.IX.2008
NHMO Lep08298	*Epirrita autumnata*	Norway	Lierne[2]	64°26′44.32″N, 13°42′57.43″E	25.IX.2008
NHMO Lep08283	*Epirrita autumnata*	Norway	Hol[2]	60°31′33.79″N, 8°18′21.14″E	5-7.IX.2008
NHMO Lep08284	*Epirrita autumnata*	Norway	Hol[2]	60°31′33.79″N, 8°18′21.14″E	5-7.IX.2008
NHMO Lep08285	*Epirrita autumnata*	Norway	Hol[2]	60°31′33.79″N, 8°18′21.14″E	5-7.IX.2008
NHMO Lep08286	*Epirrita autumnata*	Norway	Hol[2]	60°31′33.79″N, 8°18′21.14″E	5-7.IX.2008
RVO014	*Epirrita autumnata*	Norway	Nøtterøy[2]	59°12′20.38″N, 10°33′49.53″E	5.VIII.2006
KB09008	*Epirrita autumnata*	Norway	Oslo[2]	59°53′52.18″N, 10°43′55.37″E	4-23.IX.2008
KB09009	*Epirrita autumnata*	Norway	Oslo[2]	59°53′52.18″N, 10°43′55.37″E	21.VIII-4.IX.2008
NHMO Lep09111	*Epirrita autumnata*	Norway	Hobøl[2]	59°38′30.02″N, 10°59′50.52″E	25-26.IX.2009
KB09016	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	IX.2007
KBE07032	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07033	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07034	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07035	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07036	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07037	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07038	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07039	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07040	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07041	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07042	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
KBE07043	*Epirrita autumnata*	Norway	Vang[2]	61°09′29.51″N, 8°31′40.88″E	1.IX.2006
NHMO Lep07087	*Epirrita autumnata*	Norway	Ås[2]	60°04′02.28″N, 10°11′15.06″E	27.IX.2007
KB08011	*Epirrita autumnata*	Norway	Lillesand[3]	58°11′07.17″N,8°13′58.90″E	4.X.2007
KB08012	*Epirrita autumnata*	Norway	Lillesand[3]	58°11′07.17″N,8°13′58.90″E	X.2007
KB08013,5	*Epirrita autumnata*	Norway	Åseral[3]	58°47′11.51″N,7°19′12.63″E	11.X.2007
KB08092	*Epirrita autumnata*	Norway	Søgne[3]	58°07′37.47″N,7°36′40.09″E	7.X.2008
KB08093	*Epirrita autumnata*	Norway	Søgne[3]	58°07′37.47″N,7°36′40.09″E	17.X.2008
KB08094	*Epirrita autumnata*	Norway	Søgne[3]	58°07′37.47″N,7°36′40.09″E	17.X.2008
KB08095	*Epirrita autumnata*	Norway	Søgne[3]	58°07′37.47″N,7°36′40.09″E	17.X.2008
KB09097	*Epirrita autumnata*	Norway	Søgne[3]	58°07′37.47″N,7°36′40.09″E	17.X.2008
KB08083	*Epirrita autumnata*	Norway	Kristiansand[3]	58°09′39.95″N, 8°05′57.44″E	29.IX.2008
KB08086	*Epirrita autumnata*	Norway	Kristiansand[3]	58°09′39.95″N, 8°05′57.44″E	6.X.2008
KB08087	*Epirrita autumnata*	Norway	Kristiansand[3]	58°12′08.76″N, 8°06′05.92″E	12.X.2008
KB08101	*Epirrita autumnata*	Norway	Kristiansand[3]	58°12′08.76″N, 8°06′05.92″E	18.X.2008
KB09103	*Epirrita autumnata*	Norway	Kristiansand[3]	58°12′08.76″N, 8°06′05.92″E	12.X.2008
KB09004	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	5-15.IX.2008
KB09005	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	5-15.IX.2008
KB09020	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	IX.2008
KB09021	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	IX.2008
KB09023	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	IX.2008
KB09024	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	X.2008
KB09025	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	X.2008
KB09026	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	X.2008
KB09027	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	X.2008
KB09028	*Epirrita autumnata*	Norway	Årdal[4]	61°21′19.81″N,7°52′47.72″E	X.2008
RVO021	*Epirrita autumnata*	Norway	Suldal[4]	59°29′49.34″N, 6°15′30.65E″	2.X.2002
RVO010	*Epirrita autumnata*	Norway	Suldal[4]	59°32′50.25″N,6°23′13.76″E	18.X.2004

(*Continued*)

Appendix: Table A1. (Continued).

Tissue sample	Species	Country	Municipality	Coordinates	Collecting date
RVO011	*Epirrita autumnata*	Norway	Suldal[4]	59°39′27.09″N,6°52′06.80″E	24.X.2002
RVO013	*Epirrita christyi*	Norway	Klepp	58°44′23.49″N, 5°30′45.48″E	14.X.2001
RVO015	*Epirrita christyi*	Norway	Nøttery	59°12′20.38″N, 10°33′49.53″E	IX.2006
RVO016	*Epirrita christyi*	Norway	Nøttery	59°12′20.38″N, 10°33′49.53″E	IX.2006
RVO001	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
RVO002	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
RVO003	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
RVO004	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
RVO005	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
RVO007	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
RVO008	*Epirrita christyi*	Norway	Skien	59°08′00.51″N, 9°39′25.22″E	11.IX.2008
KB09033	*Epirrita christyi*	Norway	Årdal	61°21′19.81″N, 7°52′47.72″E	17.X.2008
KBE08098	*Epirrita christyi*	Norway	Kristiansand	58°09′39.95″N, 8°05′57.44″E	18.X.2008
KBE08100	*Epirrita christyi*	Norway	Kristiansand	58°12′08.76″N, 8°06′05.92″E	18.XI.2008
NHMO Lep09112	*Epirrita christyi*	Norway	Asker	59°50′10.26″N, 10°28′01.80″E	1.X.2009
NHMO Lep09113	*Epirrita christyi*	Norway	Asker	59°50′10.26″N, 10°28′01.80″E	1.X.2009
NHMO Lep09114	*Epirrita christyi*	Norway	Asker	59°50′10.26″N, 10°28′01.80″E	1.X.2009
NHMO Lep07088	*Epirrita christyi*	Norway	Ås	60°04′02.28″N, 10°11′15.06″E	27.IX.2007
KBE07046	*Epirrita christyi*	Norway	Lillesand	58°11′07.17″N,8°13′58.90″E	30.IX.2007
KB08102	*Epirrita dilutata*	Norway	Kristiansand	58°12′08.76″N, 8°06′05.92″E	28.X.2008
KBE08096	*Epirrita dilutata*	Norway	Kristiansand	58°12′08.76″N, 8°06′05.92″E	18.X.2008
KBE08097	*Epirrita dilutata*	Norway	Kristiansand	58°12′08.76″N, 8°06′05.92″E	18.X.2008
KBE08099	*Epirrita dilutata*	Norway	Kristiansand	58°12′08.76″N, 8°06′05.92″E	18.X.2008
RVO012	*Epirrita dilutata*	Norway	Kristiansand	58°04′06.92″N, 7°58′52.55″E	1.XI.1999
KBE07047	*Epirrita dilutata*	Norway	Lillesand	58°11′07.17″N,8°13′58.90″E	14.X.2007
NHMO Lep09105	*Epirrita autumnata*	Finland	Lohja	60°15′01.16″N, 24°04′45.68″E	29.IX.2008
NHMO Lep09106	*Epirrita autumnata*	Finland	Lohja	60°15′01.16″N, 24°04′45.68″E	29.IX.2008
NHMO Lep09107	*Epirrita autumnata*	Finland	Lohja	60°15′01.16″N, 24°04′45.68″E	29.IX.2008
NHMO Lep09108	*Epirrita autumnata*	Finland	Hyvinkââ	60°37′54.63″N, 24°51′51.13″E	8.X.2008
NHMO Lep09109	*Epirrita autumnata*	Finland	Lohja	60°15′01.16″N, 24°04′45.68″E	29.IX.2008
NHMO Lep09118	*Epirrita autumnata*	Scotland	Banffshire	57°25′11.28″N, 2°38′35.08″V	18.X.2009
NHMO Lep09119	*Epirrita autumnata*	Scotland	Banffshire	57°25′11.28″N, 2°38′35.08″V	18.X.2009
NHMO Lep09120	*Epirrita autumnata*	Scotland	Banffshire	57°25′11.28″N, 2°38′35.08″V	18.X.2009

[1]North.
[2]East.
[3]South.
[4]West.

Appendix: Table A2. GenBank accession numbers.

Tissue sample	Species	COI	ITS2	Wingless	*wsp*
NHMO Lep08305	*Epirrita autumnata*	JX260769	JN225585	JN225572	
NHMO Lep08306	*Epirrita autumnata*	JX260785	JN225586	JN225573	
NHMO Lep08307	*Epirrita autumnata*	JX260741	JN225587	JN225566	
NHMO Lep08308	*Epirrita autumnata*	JX260738	JN225588	JN225581	
RVO009	*Epirrita autumnata*	JX260749	JN225606	JN225568	JX310341
NHMO Lep08297	*Epirrita autumnata*	JX260733			
NHMO Lep08298	*Epirrita autumnata*	JX260786			
NHMO Lep08283	*Epirrita autumnata*	JX260759	JN225584	JN225571	
NHMO Lep08284	*Epirrita autumnata*	JX260775			
NHMO Lep08285	*Epirrita autumnata*	JX260783			
NHMO Lep08286	*Epirrita autumnata*	JX260782			

(*Continued*)

Appendix: Table A2. (Continued).

Tissue sample	Species	COI	ITS2	Wingless	*wsp*
RVO014	*Epirrita autumnata*	JX260772	JN225590	JN225575	
KB09008	*Epirrita autumnata*	JX260758	JN22600		
KB09009	*Epirrita autumnata*	JX260768	JN225601		
NHMO Lep09111	*Epirrita autumnata*	JX260754			
KB09016	*Epirrita autumnata*	JX260757			
KBE07032	*Epirrita autumnata*	JX260784			
KBE07033	*Epirrita autumnata*	JX260739			
KBE07034	*Epirrita autumnata*	JX260760			
KBE07035	*Epirrita autumnata*	JX260762			
KBE07036	*Epirrita autumnata*	JX260752			
KBE07037	*Epirrita autumnata*	JX260771			
KBE07038	*Epirrita autumnata*	JX260773			
KBE07039	*Epirrita autumnata*	JX260789			
KBE07040	*Epirrita autumnata*	JX260788			
KBE07041	*Epirrita autumnata*	JX260756			
KBE07042	*Epirrita autumnata*	JX260743			
KBE07043	*Epirrita autumnata*	JX260753			
NHMO Lep07087	*Epirrita autumnata*	JX260779			
KB08011	*Epirrita autumnata*	JX269737	JN225592	JN225580	
KB08012	*Epirrita autumnata*	JX260787			JX310345
KB08013,5	*Epirrita autumnata*	JX260778			
KB08092	*Epirrita autumnata*	JX260745			
KB08093	*Epirrita autumnata*	JX260736	JN225595		JX310346
KB08094	*Epirrita autumnata*	JX260744	JN225596	JN225567	
KB08095	*Epirrita autumnata*	JX260742			
KB09097	*Epirrita autumnata*	JX260780	JN225604		
KB08083	*Epirrita autumnata*	JX260770			
KB08086	*Epirrita autumnata*	JX260751			
KB08087	*Epirrita autumnata*	JX260791	JN225593	JN225577	JX310342
KB08101	*Epirrita autumnata*	JX260746			
KB09103	*Epirrita autumnata*	JX260740	JN225605	JN225569	JX310340
KB09004	*Epirrita autumnata*	JX260731			
KB09005	*Epirrita autumnata*	JX260761			
KB09020	*Epirrita autumnata*	JX260790			
KB09021	*Epirrita autumnata*	JX260765			
KB09023	*Epirrita autumnata*	JX260767			
KB09024	*Epirrita autumnata*	JX260766			
KB09025	*Epirrita autumnata*	JX260792			
KB09026	*Epirrita autumnata*	JX260774	JN225602	JN225578	
KB09027	*Epirrita autumnata*	JX260734	JN225603		
KB09028	*Epirrita autumnata*	JX260732			
RVO021	*Epirrita autumnata*	JX260747	JN225591	JN225576	
RVO010	*Epirrita autumnata*	JX260710	JN225589	JN225574	JX310348
RVO011	*Epirrita autumnata*	JX260781			
RVO013	*Epirrita christyi*	JX260808			
RVO015	*Epirrita christyi*	JX260796			
RVO016	*Epirrita christyi*	JX260802			
RVO001	*Epirrita christyi*	JX260800			
RVO002	*Epirrita christyi*	JX260806			
RVO003	*Epirrita christyi*	JX260794			
RVO004	*Epirrita christyi*	JX260807	JN225607	JN225582	
RVO005	*Epirrita christyi*	JX260793	JN225608	JN225583	
RVO007	*Epirrita christyi*	JX260799			JX310336
RVO008	*Epirrita christyi*	JX260797			
KB09033	*Epirrita christyi*	JX260804			

(*Continued*)

Appendix: Table A2. (Continued).

Tissue sample	Species	COI	ITS2	Wingless	*wsp*
KBE08098	*Epirrita christyi*	JX260809			
KBE08100	*Epirrita christyi*	JX260805			JX310335
NHMO Lep09112	*Epirrita christyi*	JX260803			JX310337
NHMO Lep09113	*Epirrita christyi*	JX260810			JX310338
NHMO Lep09114	*Epirrita christyi*	JX260795			JX310339
NHMO Lep07088	*Epirrita christyi*	JX260798			
KBE07046	*Epirrita christyi*	JX260801			
KB08102	*Epirrita dilutata*	JX260814	JN225610		
KBE08096	*Epirrita dilutata*	JX260813			
KBE08097	*Epirrita dilutata*	JX260816	JN225609		
KBE08099	*Epirrita dilutata*	JX260812			
RVO012	*Epirrita dilutata*	JX260811			
KBE07047	*Epirrita dilutata*	JX260815			
NHMO Lep09105	*Epirrita autumnata*	JX260764	JN225594	JN225570	JX310347
NHMO Lep09106	*Epirrita autumnata*	JX260735			
NHMO Lep09107	*Epirrita autumnata*	JX260763			
NHMO Lep09108	*Epirrita autumnata*	JX260748			JX310343
NHMO Lep09109	*Epirrita autumnata*	JX260777			JX310344
NHMO Lep09118	*Epirrita autumnata*	JX260750	JN225598		
NHMO Lep09119	*Epirrita autumnata*		JN225599	JN225579	
NHMO Lep09120	*Epirrita autumnata*	JX260776	JN225597		

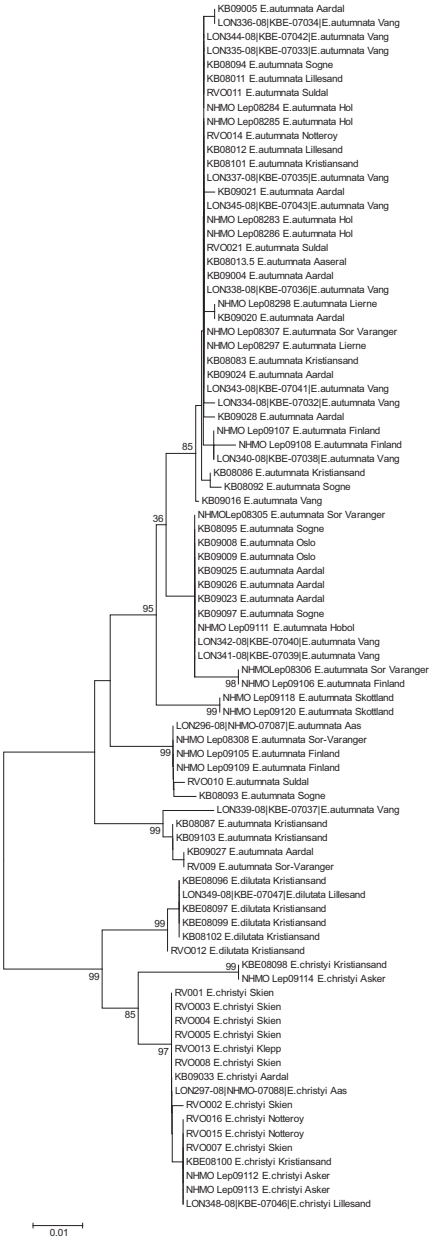


Figure A1. Minimum evolution analysis (Tamura 3-parameter as substitution model) of the COI data set (n = 86), showing the same structure as the NJ analysis, with 5 distinct COI-haplotypes. Bootstrap (10000 iterations) is shown at each node.

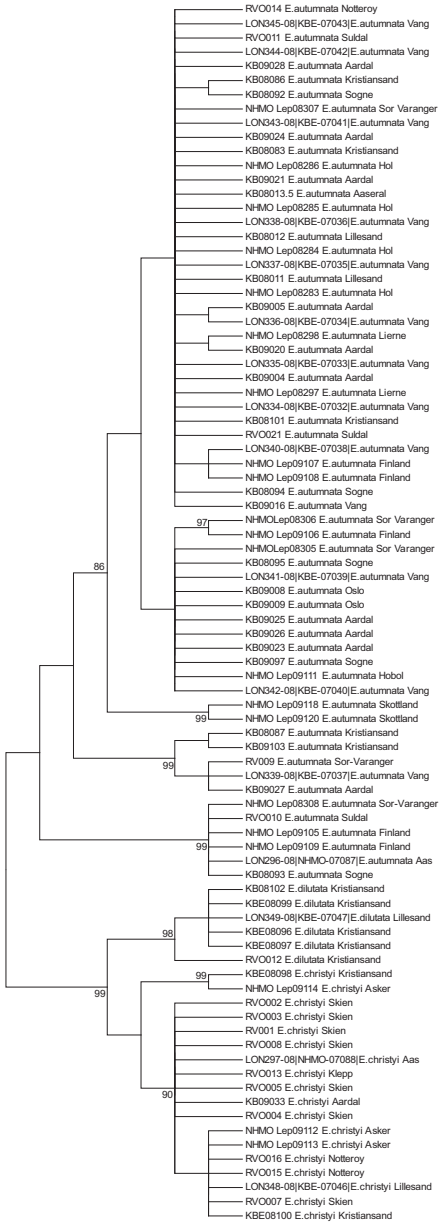


Figure A2. Maximum parsimony analysis (using the complete deletion option) of the COI data set (n = 86). MP analysis also shows high intraspecific variation within E. autumnata. Bootstrap values (500 iterations) are shown at each node.

III

Rapid sperm evolution in the bluethroat (*Luscinia svecica*) subspecies complex.

Silje Hogner[1], Terje Laskemoen[1], Jan T. Lifjeld[1], Václav Pavel[2], Bohumír Chutný[3], Javier García[4], Marie-Christine Eybert[5], Ekaterina Matsyna[6], and Arild Johnsen[1].

[1]Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, NO-0318 Oslo, Norway

[2]Laboratory of Ornithology, Palacký University, Svobody 26, 771 46 Olomouc, Czech Republic
[3]Malinová 27, 106 00 Praha 10, Czech Republic

[4] Área de Ecología, Facultad de Ciencias Biológicas y Ambientales, Universidad de León, Campus de Vegazana s/n, 24071 León, Spain
[5] University of Rennes 1, UMR 6553, 263 Avenue du Général Leclerc, CS 74205, 35042 Rennes cedex, France
[6]Laboratory of Ornithology, Ecological Centre "Dront", Nizhny Novgorod, 603000, P.O. Box 631, Russia

Correspondence:

Silje Hogner

E-mail: silje.hogner@nhm.uio.no

Phone: +47 22 85 18 86

Fax: +47 22 85 18 37

Running title: Sperm evolution in bluethroats

Abstract

Spermatozoa are among the most variable animal cell types, and much research is currently directed towards explaining inter- and intraspecific variation in sperm form and function. Recent comparative studies in passerine birds have found associations between the level of sperm competition and both sperm length and sperm velocity. In species with sperm competition, postcopulatory sexual selection may shape the morphology of sperm as adaptations to the female environment. The speed of evolutionary change in sperm morphology at the species level is largely unknown. In this study, we analysed variation in sperm morphology among morphologically distinct and geographically isolated bluethroat subspecies in Europe. Consistent with previous studies, our analyses of mtDNA and nuclear introns suggest recent divergence and lack of lineage sorting among the subspecies. We found significant divergence in total sperm length and in the length of some sperm components (i.e. head and midpiece). There was a significantly positive relationship between pairwise divergences in sperm morphology and mitochondrial DNA, suggesting a role for genetic drift in sperm divergence. The magnitude of sperm length divergence was considerably higher than that in other geographically structured passerines, and even higher than that observed between several sister species. We hypothesize that the rapid sperm evolution in bluethroats is driven by sperm competition, and that strong postcopulatory sexual selection on sperm traits can lead to rapid speciation through reproductive incompatibilities.

Keywords: sperm competition, sperm morphology, sperm size variation, reproductive isolation

Introduction

Sexual selection is a potent evolutionary force that may lead to rapid changes in traits related to success in competition for mates, like ornaments and armaments (Andersson 1994). Divergence in such secondary sexual traits and associated mate preferences may lead to precopulatory prezygotic isolation between populations and hence catalyse the early stages of speciation (Panhuis et al. 2001; Grethner 2010; Maan and Seehausen 2011). Similarly, postcopulatory sexual selection (i.e. sperm competition and cryptic female choice) exerts strong selection on primary sexual traits (e.g. gametes and the tissues producing them), and divergence in such traits may also lead to prezygotic reproductive isolation, and ultimately speciation (Coyne and Orr 2004).

Spermatozoa are among the most variable animal cell types, exhibiting considerable variation in morphometry and behaviour at all levels of organisation (e.g. Cohen 1977; Pitnick et al. 2009). Recent comparative studies have suggested that sperm competition influences the evolution of a range of male reproductive traits (Birkhead and Møller 1998). For example, species experiencing stronger sperm competition have larger testes (relative to their body size) (Møller 1991; Harcourt et al. 1995; Hosken 1997; Stockley et al. 1997; Byrne et al. 2002) and produce more sperm (Rowe and Pruett-Jones 2011). Furthermore, total sperm length is associated with the risk of sperm competition in a range of taxa including, insects (Gage 1994; Morrow and Gage 2000), fish (Stockley et al. 1997; Balshine et al. 2001), frogs (Byrne et al. 2003), mammals (Gomendio and Roldan 1991; Breed and Taylor 2000; but see Hosken 1997; Gage and Freckleton 2003) and birds (Lüpold et al. 2009b; Kleven et al. 2009; but see Immler and Birkhead 2007).

Variation in sperm size between males in a population is also associated with sperm competition. Specifically, the coefficient of between-male variation in total sperm length is negatively related to the level of sperm competition faced by males in passerine birds (Calhim et al. 2007; Kleven et al. 2008; Lifjeld et al. 2010), suggesting that sperm competition exerts strong stabilizing selection on sperm morphology. Sperm traits may also respond to directional selection from sperm competition (e.g. Kleven et al. 2009; Lüpold et al. 2009a), in addition to being influenced by drift (Laskemoen et al. in press). Moreover, sperm morphology appears to be shaped via interaction with the female reproductive environment (Woolley 1970; Briskie et al. 1997; Morrow and Gage 2001; Higginson et al. 2012). Therefore, a complex range of selection pressures and drift are likely to drive the evolution of sperm morphology and, consequently, sperm traits may

move towards different optima in different populations, especially if the spatially isolated populations differ in strength or form of selection.

Passerine birds exhibit considerable variation in total sperm length, ranging from 42.7µm in the red-backed shrike (*Lanius collurio*) (Briskie et al. 1997) to 291µm in reed bunting (*Emberiza schoeniclus*) (Dixon and Birkhead 1997). Sperm length generally shows high phylogenetic dependence in comparative studies (e.g. Kleven et al. 2009). However, closely related species may also exhibit large divergence in total sperm length, e.g. sand martins *Riparia riparia* (123.0 µm) (Kleven et al. 2009) and tree swallows *Tachycineta bicolor* (235.4 µm) Laskemoen et al. 2010). Within the family Muscicapidae, total sperm length varies from 101.2 µm (collared flycatcher *Ficedula albicollis*) to 279.9 µm (nightingale *Luscinia megarhynchos*) among 12 investigated species (own unpublished data). To date, relatively few studies have investigated intraspecific variation in sperm morphology, despite the fact that analyses at these levels (between-individuals within a population and between populations) should provide considerable insight into both the selection pressure shaping the evolution of sperm morphology and the speed of evolutionary change in sperm traits. Additionally, intraspecific studies may reveal contrasting patterns to studies at the interspecific level (e.g. Lüpold et al. 2009b), thus a comprehensive understanding of the evolution of sperm traits requires examination at both levels of organisation.

A limited number of recent studies of passerine birds demonstrate significant variation in sperm morphology among different populations (Schmoll and Kleven 2011; Laskemoen et al. in press; Lüpold et al. 2011). Moreover, in addition to variation in sperm morphology among redwing-blackbird (*Agelaius phoeniceus*) populations, Lüpold et al. (2011) found a gradual increase in sperm size from southwest to northeast of the breeding range, and a negative relationship between sperm length and body size. However, in this study and the study of Schmoll and Kleven (2011) on coal tits (*Periparus ater*), the degree of genetic differentiation between subspecies was unknown and thus no inference could be drawn regarding the speed of sperm diversification. Laskemoen et al. (in press) found significant variation in sperm morphology among barn swallow populations (*Hirundo rustica*). Moreover, in that study, the subspecies with highest genetic distance also showed more differences on sperm morphology, leading the authors to hypothesise that variation in sperm morphology between subspecies might reflect the genetic distance between taxa (Laskemoen et al. in press). Studies incorporating information on both variation in sperm

traits along with data concerning genetic divergence are necessary to examine the speed and direction of evolution on sperm traits.

Here, we examined both sperm morphology and genetic divergence in the bluethroat (*Luscinia svecica*) subspecies complex. The bluethroat is a small (~18 g) passerine bird, which ranges from upper arctic limits to temperate and steppe middle latitudes and breeds from the western Palearctic to eastern Eurasia (Cramp 1988). Currently, ten subspecies are recognized based primarily on differences in male plumage characteristics and to some extent size (Cramp 1988). All subspecies are sexually dimorphic: males are larger and exhibit striking sexual ornamentation, whereas female generally exhibit drab plumage (Johnsen et al. 2006). Specifically, males have a colourful throat patch with blue and chestnut surrounding a conspicuous central spot (white, red or absent depending on the subspecies), which is displayed during courtship (Peiponen 1960; Johnsen and Lifjeld 1995). There is also evidence that subspecies vary in song characteristics (Turcokova et al. 2010). Currently, the phylogenetic relationships among the bluethroat subspecies are not well resolved and subspecific status is somewhat contentious. Using mtDNA markers (control region, *cyt b*), Questiau et al. (1998) found support for two distinct subspecies clusters (*svecica* and *namnetum*), whereas Zink et al. (2003), despite reporting a relatively high degree of population differentiation (F_{ST} = 0.29), found little support for subspecies recognition in a study of seven morphs previously identified as subspecies based on morphological differences. Using 11 microsatellite loci, Johnsen et al. (2006) found evidence for differentiation (F_{ST} = 0.042) across bluethroat populations in Europe and Asia. Moreover, that study found support for genetic differentiation between some morphologically distinct subspecies, most notably *svecica, namnetum, azuricollis* and *cyanecula* (Johnsen et al. 2006).

Laskemoen *et al*. (2007) demonstrated considerable between-male and within-male variation in sperm morphology in the nominate subspecies. In the present study, we focus on the four most distinct subspecies identified by Johnsen et al. (2006) (i.e. ssp. *svecica, namnetum, azuricollis, cyanecula* and in addition ssp. *volgae*), and investigate between-population variation in sperm morphology in relation to variation in two mtDNA regions and two nuclear introns. Our study had two main aims: 1) to test if there is a relationship between genetic divergence and sperm divergence within these five study populations and 2) to compare differences in sperm evolution between bluethroat and other species with known sperm divergence.

Material and methods

Field work

We collected samples from each of the five bluethroat subspecies (*azuricollis, cyanecula, namnetum, svecica and volgae*) (see Table S1 for details of localities and Table S2 for information about individual samples) during the peak of the breeding season in 1996-2011. Birds were caught on their home territories using song playback and mist nets or clap nets using mealworm as bait. Sperm samples were collected via cloacal massage (Wolfson 1952). The ejaculate was collected with a micro capillary tube and fixed in 5% formalin (PBS) solution. At the same time, 25 µl of blood was collected by brachial venipuncture and stored in 96% ethanol. DNA was later extracted following the protocol for the E.Z.N.A blood kit (Omega Bio-Tek, Inc, Norcross, Georgia) and the QIAmp blood kit (QIAGEN, Inc. Valencia, California).

Sperm morphology

For each individual, approximately 15µl of diluted sperm was applied to a glass microscope slide and allowed to air-dry. Next, the slide was gently rinsed with distilled water and allowed to air-dry once again. We captured digital images (160 x magnification) of sperm cells using a Leica DFC420 camera mounted on a Leica DM6000 B digital light microscope (Leica Microsystems, Switzerland) and measured sperm length using specialised image analysis software (Leica Application suite v. 2.6.0 R1). For each male, the following measurements were obtained (±0.1 µm): head length, midpiece length, tail length and total length (i.e. head + midpiece + tail). Following the recommendation of Laskemoen et al. (2007) we measured 10 morphologically normal sperm from a minimum of 10 males per subspecies (with the exception of *volgae*, for which samples were only available from nine males). Additionally, we calculated the within male (CV_{wm}) and between male coefficient of variation in total sperm length (CV_{bm}), the latter being an index for sperm competition across passerines (Lifjeld et al. 2010). We then adjusted CV_{bm} according to the sample size using the formula $CV_{bm} + (1/(4n))$ (Sokal & Rohlf, 1995), because CV_{bm} has been documented to be underestimated in small population samples (Laskemoen et al., 2007). Hereafter, CV_{bm} will refer to the adjusted value. To avoid observer effects, all measurements were conducted by one person (TL).

PCR and sequencing

We sequenced two mitochondrial regions (COI-region and control region, n = 84 each) and two Z-linked introns (BRM-15 and VLDLR-7, n=53 each; see Table S3 for primer combinations and PCR conditions). All markers were amplified in PCR reaction volumes of 10 µL, containing dH_2O, 1X PCR buffer II (Applied Biosystems), 1.5mM magnesium, 0.2mM dNTP (ABgene, Epsom, UK), 0.5mM forward and reverse primer, 3% dimethyl sulfoxide (DMSO), 0.25 U AmpliTaq DNA polymerase (Applied Biosystems, California) and approximately 50ng DNA template. Amplifications were run on a DNA Engine Tetrad 2 (MJ Research, Waterton, MA, USA) using the following profile: 95˚C for 1 min, 94˚C for 30 sec, annealing temperature 52-56˚C (depending on primer combination; see Table S3) for 30 sec, 72˚C for 1 min. This profile was then repeated for a further 34 cycles before a final elongation step of 72˚C for 10 min. In order to confirm amplification success and to exclude any contamination, 3 µl of PCR-product was electrophoresed in 1% agarose TBE.

The remaining PCR product was purified by digesting unincorporated nucleotides and primers using diluted (1:9) ExoSap-It (United States Biochemical, Cleveland) run at 37^0C for 45 min, followed by 15min at 80^0C to inactivate the enzyme. The PCR products were then sequenced using BigDye Terminator v 3.1 Cycle Sequencing kit (Applied Biosystems, California). Next, sequences were aligned (using ClustalW) and manually edited in the program Mega 5.0 (Tamura et al. 2007). For each locus, all sequences were truncated to the length of the shortest sequence for comparison. Altogether, we included samples from five different subspecies and seven different populations. Sperm samples and mitochondrial regions were analysed from individuals sampled at the same locality during a single collection episode, with the exception of individuals sampled in Wroclaw (Poland) for which blood and sperm were collected from the same locality but from different birds during different years (2009 and 2010, respectively). For the Z-introns, samples were obtained from the same localities, but in different years, as both sperm and mtDNA samples, with the exception of *cyanecula,* for which samples were obtained from Thüringen (Germany) and Trebon (Czech Republic). Importantly, introns were sequenced from females only which allowed us to obtain exact haplotypes, since female birds are hemizygous on the Z-chromosome. Finally, we did not sequence any introns from the *volgae* population.

Statistical methods

We tested for differences in sperm morphology (means per male) among subspecies using ANOVAs and *posthoc* Tukey HSD tests using Statistica v 7.1 (StatSoft Inc). DNAsp v 5 (Librado and Rozas 2009) was used to calculate the nucleotide diversity (π) (Hudson *et al.* 1987) and Tajima's *D* (Tajima 1989) for both mitochondrial and nuclear regions. We implemented a model test in Mega v5.05 (Tamura et al. 2011), using BIC (Bayesian Information Criterion) scores to find the best fitting substitution model for the two mitochondrial markers combined (all further analyses were conducted for the two mtDNA regions combined) and each intron separately. For mitochondrial regions, a gene tree based on maximum likelihood was made using Mega v5.05 with the substitution model HKY+G, with *Luscinia megarhynchos* as outgroup. For nuclear regions, maximum likelihood were constructed using the K2P (VLDLR-7) and Tamura-3-parameter (BRM-15) model using the same outgroup. Bootstrap values were calculated in Mega v5.05 using 10 000 iterations. Translation from nucleotide to amino acid sequences of the analysed regions revealed no stop codons, frameshifts or systematic double peaks in the COI region, indicating an absence of pseudogenes.

In order to examine the genetic structure of the populations, analysis of molecular variance (AMOVA) was run using Arlequin v3.5 (Excoffier et al. 2005). The variance was partitioned into variation between populations and variation within populations. Pairwise species differentiation was estimated using F_{ST} (Weir and Cockerham 1984), with default settings in the population comparison. These F_{ST} values can be used as short term genetic distances between populations. The null hypothesis is no differences between the populations, and the P-value is given as the proportions of simulations giving a F_{ST} value larger or equal to the observed one. Similarly, we calculated a measure of phenotypic divergence, P_{ST}, that expresses the proportion of total variance in sperm length that can be attributed to the variation among populations. P_{ST} was calculated using the following formula, $P_{ST} = \frac{\sigma_B^2}{c\sigma_B^2 + 2h^2\sigma_W^2}$, where σ_B^2 is the phenotypic variance between populations, σ_W^2 is the phenotypic variance within populations and h^2 denotes the heritability (Leinonen et al. 2006). We calculated σ_B^2 and σ_W^2 from one-factor ANOVA following Sokal and Rohlf (1995; p. 216, Box 9.2). The heritability was conservatively set to 0.62, based on an estimate of heritability of flagellum length (midpiece plus tail) reported for the zebra finch (*Taeniopygia guttata*) by Birkhead et al. (2005). Next, in order to test for a correlation between pairwise estimates of genetic divergence (F_{ST}) and sperm length divergence (P_{ST}),

we performed a Mantel test (Mantel 1967) to correct for the multiple uses of the same populations. The Mantel test was conducted in R version 2.11.1 (R Development Core Team, 2007) using the package ade4 (Dray and Dufour 2007) and the P-value obtained through Monte Carlo simulations with 9999 replicates.

Results

Sperm morphology

We found significant variation in total sperm length among the five subspecies (Table 1) ($F_{4,77} = 47.0$, $P < 0.001$; Fig. 1). This was more or less explained by variation in midpiece length ($F_{4,77} = 26.8$, $P < 0.001$) which showed the same pattern as total length. In addition, we found significant variation in head length ($F_{4,77} = 7.47$, $P < 0.001$), although this component showed a different pattern than total and midpiece length. In contrast, tail length did not vary between subspecies ($F_{4,77} = 0.80$, $P = 0.53$). Posthoc tests revealed that total length and midpiece length was significantly longer in *azuricollis* relative to all other subspecies (Table 2). In addition, *svecica* had significantly longer total sperm length relative to *cyanecula*, and *cyanecula* had a significantly longer midpiece relative to *namnetum* (Table 2). Finally, *svecica* had significantly longer sperm head length relative to both *cyanecula* and *namnetum* (Table 2). The total P_{ST} was 0.70, $P < 0.001$.

We found a nearly two-fold variation in CV_{bm} of total sperm length, ranging from 1.91 in *svecica* to 3.30 in *azuricollis* (Table 1). A post hoc test of homogeneity of variances showed significant heterogeneity of sperm variation among populations (Bartletts test $\chi^2 = 9.59$, $P = 0.048$). Moreover, the CV_{wm} varied significantly between populations ($F_{4,77} = 8.44$, $P < 0.001$) and ranged from 1.30 to 1.92 (Table 1), with *cyanecula* differing significantly from both *azuricollis* and *svecica* (Table 2). There was no significant correlations between CV_{bm} and mean CV_{wm} (Spearman's correlation: $n = 5$, $r_s = 0.05$, $P = 0.93$).

Variation in mitochondrial DNA

Of the 1229 bp sequenced for mtDNA in all individuals, we found 27 variable sites and 16 parsimony informative sites (*Luscinia megarhynchos* outgroup excluded). The nucleotide diversity (π) for all subspecies combined was 0.0023, and ranged from 0.0006 (*azuricollis*) to 0.0022 (*volgae*) across the five subspecies (Table S4). A neutrality test based on allele

frequency, Tajima's D, was performed for the two mitochondrial regions combined, for each of the subspecies and none of these test showed significant results (Table S4).

The mitochondrial gene tree did not show a clear structure of the subspecies in monophyletic clades, but it did show moderate support (with bootstrap support just above 50%) for two groups, one consisting of *namnetum*, *cyanecula* and *azuricollis*, and the other consisting of *svecica* (Fig 2). Interestingly, the *volgae* population was present in both groups (Fig 2). In addition, the *azuricollis* population formed a monophyletic group, with relatively high bootstrap support (> 80%) for a clade consisting of 19 out of the 21 individuals from this population. The mtDNA gene trees did not have enough resolution or support to determine the most ancestral of these five subspecies.

AMOVA revealed that 61% of the total variation occurred between subspecies while 39% occurred within populations, with an overall significant F_{ST}-value ($F_{ST} = 0.61$, $P < 0.001$). Additionally, pairwise F_{ST} values between the five subspecies showed relatively high and significant values (see Table 3). The maximum genetic distance between any individual (using K2P substitution model) was estimated to 0.7% and the mean genetic distance to 0.3%. Assuming a conventional molecular clock and mutation rate (i.e. 2% divergence per million years, Bromham and Penny 2003), the maximum time since divergence was 350 000 ya and the mean time since divergence150 000 ya.

Variation in nuclear DNA

Of the 933bp sequenced for the nuclear introns in all individuals, we found 39 variable sites and 22 parsimony informative sites (outgroup excluded). The total nucleotide diversity (π) for the introns was 0.0061 for all subspecies combined. The π-value ranged from 0.0006 (in *azuricollis*) to 0.0074 (in *cyanecula*) for the two introns combined (Table S4). The neutrality test, Tajima's D, showed no significant values for any of the subspecies using introns (Table S4). The two gene trees calculated using the introns BRM-15 and VLDLR-7 (with *Luscinia megarhynchos* as outgroup) showed no structure related to the subspecies (Figs. S1 and 2). AMOVA revealed that 12% of the total variation occurred between populations, while 88% occurred within populations, with an overall significant F_{ST}-value ($F_{ST} = 0.12$, *P*-value < 0.001).

Comparing sperm with genes

There was an overall good concordance between genetic divergence ($F_{ST} = 0.61$) and sperm length divergence ($P_{ST}= 0.70$) for all five populations analysed together. There was

also a significant positive correlation between the two divergence measures for all pairwise estimates among populations (Figure 3; Mantel test: $R = 0.80$, $P = 0.042$). There was, however, no such relationship between intronic F_{ST}-estimates and P_{ST}-estimates (Mantel test: $R = 0.24$, $P = 0.34$).

Populations with high genetic diversity in mtDNA and Z introns had significantly lower variation in sperm length (mtDNA $F_{ST} = 0.61$, intronic $F_{ST} = 0.12$ versus total $P_{ST} = 0.70$). This is opposite to expectation from a neutral model of variation in a heritable character, like sperm length.

Discussion

Corresponding divergence in the bluethroats

We found considerable variation in sperm morphology between subspecies, with differences in sperm head and midpiece length, as well as total sperm length between several of the sampled taxa. Seventy percent of the phenotypic variation in total sperm length was confined to the between-population level, while 60% of the mitochondrial DNA and 12% of the nuclear variation resided between populations. The maximum estimated time since divergence among individuals in our sample was estimated to be 350 000 ya, assuming a 2% sequence divergence per million years (Bromham and Penny 2003; Päckert et al. 2007; Weir and Schluter 2008; but see Pulquério and Nichols 2007). While we acknowledge that this estimate of time since divergence is rather crude, it is consistent with earlier suggestions that bluethroat subspecies diverged during the late Pleistocene glaciation periods (Questiau et al., 1998; Zink et al., 2003). Both mtDNA and nuclear DNA showed a lack of lineage sorting with respect to subspecies. Importantly, these findings suggest that bluethroat subspecies have diverged relatively recently and, as such, constitute 'young' taxa.

Rapid sperm evolution in the bluethroats

The high divergence in sperm morphology together with low genetic divergence suggests rapid evolution of sperm morphology in the bluethroat. Other studies have also reported inter-population variation in sperm morphology (Lüpold et al. 2011; Laskemoen et al. in press; Schmoll and Kleven 2011). However, relative divergence in total sperm length was considerably higher in bluethroats (11.6%) than between populations/subspecies in five

other species (mean divergence: 2.26%, range 0.3-3.7%; see Table 4), as well as between four pairs of sister species (mean divergence: 3.48%, range 0.3-9.9%; see Table 5).

Taken together, our findings demonstrate that the bluethroat represents a suite of young taxa exhibiting high divergence in sperm morphology. Few studies including both genetic and sperm divergence are available for comparisons. However, studies of the barn swallow suggest that European and North American populations diverged approximately 840 000 ya (1.68% divergence in CO1; Johnsen et al. 2010), but that sperm show just 3.7% relative divergence in total sperm length (Laskemoen et al. in press, Table 4). Similar comparisons can be made if we consider divergence between sister species. The relative sperm divergence between sister species ranged from 0.3% between the house sparrow (*Passer domesticus*) and the Spanish sparrow (*P. hispaniolensis,* with COI divergence = 3.0%) to 9.9% between the common redstart (*Phoenicurus phoenicurus*) and the black redstart (*P. ochrorus,* with COI divergence = 7.5%). Thus, it appears that the large divergences in sperm morphology within the bluethroat have arisen over a very short time period which suggests that evolutionary change in sperm morphology has been relatively rapid in this subspecies complex. Support for the hypothesis that sperm morphology can experience rapid evolution, has also been found in other taxa. For example, Landry et al. (2003) showed that sperm morphology evolved rapidly within two clades of the sea urchin (*Echinometra oblonga*). A genetic distance based on COI was used to estimate the time since divergence between these two clades, and suggested a split about 250 000 years ago, which is comparable in time to the mean genetic divergence found within the bluethroats (~ 150 000ya).

Rapid sperm evolution: drift versus selection?

Laskemoen et al. (in press) hypothesised that variation in sperm morphology at the subspecies level might reflect genetic distances between populations, suggesting that drift is a main driving force behind this relationship. In agreement with this, we found a significant relationship between genetic divergence and divergence in total sperm length among bluethroat subspecies (Fig. 3), suggesting that some of the variation in total sperm length can be explained by genetic drift among isolated populations. The lack of a positive relationship between neutral genetic variation and sperm variation within populations, suggests that sperm variation is not a simple function of overall genetic variation in the population. Instead, sperm variation in a population may be shaped by postcopulatory

sexual selection, mediated by the risk of sperm competition (Kleven et al. 2008, Lifjeld et al. 2010, Laskemoen et al. 2013).

Selection imposed through sperm competition is thought to be an important driver of evolutionary change in sperm traits (reviewed in Pizzari and Parker 2009). For example, because sperm optimised to the fertilising environment of the female reproductive tract are likely to experience a fertilisation advantage, it has been suggested that sperm morphology should be under stabilising selection such that all sperm match this optimal form. Moreover, as the level of sperm competition faced by males increases, the strength of stabilising selection is also expected to increase. This idea is supported by recent studies showing a negative relationship between sperm competition and variation in sperm length (Calhim et al. 2009; Kleven et al. 2009). Consequently, stabilising selection is thought to result in a reduction in sperm size variation, though such a reduction in variation does not necessarily involve a simultaneous change in mean trait values. However, there is a possibility of long-term directional evolution under a scenario of stabilizing selection given that the optimal phenotype is consistently larger or smaller than the population mean, even if the difference is minute. Hence, directional evolution does not necessarily result from directional selection, which favours phenotypes at the extreme end of a distribution. In the context of the current study, the observed pattern of divergence in sperm morphology might be explained by stabilising selection, if a reduction in variation within-populations allowed for variation between populations to be revealed.

On the other hand, sperm competition has also been associated with directional selection for sperm morphology. For example, a range of comparative and intraspecific studies have shown that sperm competition can select for both longer (Gomendio and Roldan 1991; Balshine et al. 2001; Kleven et al. 2009; Lüpold et al. 2009b) and shorter (Garcia- Stockley et al. 1997; García-González and Simmons 2007) sperm. Importantly, directional selection for sperm morphology may also explain the patterns observed here. Specifically, under this evolutionary scenario we would expect to observe divergence in mean trait values among subspecies, a pattern consistent with the results of the current study. Regardless of whether the observed pattern is determined by stabilising or directional selection, our findings suggest that sperm competition can be an important factor leading to intraspecific variation in sperm morphology.

The level of sperm competition faced by males in a population can be estimated from the between-male coefficient of variation (CV_{bm}), which is lower in species/populations exhibiting higher rates of EPP and hence higher levels of sperm

competition (Calhim et al. 2007; Kleven et al. 2008; Lifjeld et al. 2010; Hogner et al., in press; Laskemoen et al. in press), presumably because of stabilizing selection on sperm size (Lifjeld et al. 2010). We found marginally significant differences in CVbm between the populations, suggesting that they vary in the level of sperm competition, with *svecica* exhibiting the highest level (CVbm = 1.9, translating to 27% of nestlings being extrapair young according to Lifjeld et al. 2010) and *azuricollis* the lowest level (CVbm = 3.3, translating to 10% of nestlings according to Lifjeld et al. 2010). Similarly, we found significant variation in the within-male coefficient of variation (CV_{wm}), but a lack of correspondence between CV_{wm} and CV_{bm} values. This could be explained by seasonal variation in CV_{wm}, which has been shown for house wrens (*Troglodytes aedon*) (Cramer et al. in press). Variation in sperm lengths within an ejaculate is presumably not a genetically coded trait, since all sperm originate from the same diploid set of genes, but may be influenced by phenotypic plasticity or production errors during spermatogenesis.

To date, EPP data are only available for two populations. In the *svecica* population, Johnsen and Lifjeld (2003) found that 49.5% of all nests contained at least one extrapair young and 26.3% of all young were sired by extrapair males, while the corresponding figures for *namnetum* were 63.8% of nests and 41.9% of nestlings (Questiau et al. 1999), both of which are comparatively high rates of sperm competition (mean for all passerines: 27.1% of nests, 14.9% of nestlings; Griffith et al. 2002). Given the lack of data on EPP levels and other ecologically relevant variables (e.g. population density, breeding synchrony) for several of the subspecies, it is currently not possible to evaluate whether the different bluethroat subspecies are subjected to different sexual selection pressures (e.g. via differential sperm competition intensity). Nonetheless, we suggest that further investigation into differential selection pressures faced by populations/subspecies will help to unravel the relative roles of drift and sexual selection in driving sperm evolution in this system.

Directionality of evolutionary change

In addition to improving our understanding of the speed of evolutionary changes in sperm traits, our data offer some insight into the directionality of evolutionary change in sperm length. Specifically, the genetic analyses in both this study and the study by Johnsen *et al*. (2006), indicates a closer relationship among the *cyanecula*, *namnetum* and *azuricollis* subspecies, than among any of these subspecies and *svecica*. Assuming these results accurately reflect the relationship among bluethroat species, our data on total sperm length

suggest a lack of overall directionality in trait evolution. Instead, under the assumption that *azuricollis* is the ancestral subspecies (cf. Johnsen et al. 2006), then sperm length appears to have first evolved towards shorter lengths in *namnetum* and *cyanecula*, with subsequent evolution towards longer total sperm length in *svecica*. Unfortunately, however, robust testing of the directionality of trait evolution requires a relatively well-resolved phylogeny. Given that currently genetic coverage of these subspecies is relatively low (Zink et al. 2003; Johnsen et al. 2006), determining the ancestral and derived populations with the subspecies complex is not possible. Consequently, we recommend that future studies use more comprehensive genetic sampling in order to obtain a robust phylogeny and allow more sophisticated tests of directional evolution in sperm traits to be performed.

Taxonomic implications

Our results, both in terms of mtDNA and sperm morphology variation, are consistent with the findings by Johnsen et al. (2006) showing significant differentiation between the four bluethroat taxa, *svecica*, *namnetum*, *cyanecula* and *azuricollis*, thus supporting their status as independently evolving entities or taxa. This contrasts with the results of Zink et al. (2003), who found no support for subspecies using the mitochondrial control region and cyt b. It should be noted that we intentionally selected the most distinct subspecies, both based on microsatellite and phenotypic divergence (Johnsen et al. 2006) and hence the larger degree of population differentiation in mtDNA ($F_{ST} = 0.61$) compared to Zink et al. (2003) was to be expected. There are also similarities between Zink et al. (2003) ($F_{ST} = 0.29$) and our study: both studies found support for one northern and one southern group, and mitochondrial nucleotide diversity was equivalent in both studies (i.e. 0.0023). The subspecies *azuricollis*, was described by Mayaud (1958), and was only recently recognised in The Clements Checklist (Clements et al. 2012), whereas the IOC World Bird List (Gill and Donsker 2012) and leading bird handbooks (Cramp1988; del Hoyo et al. 2005) place the Iberian bluethroats in the subspecies *cyanecula*. Our data strongly support *azuricollis* as a separate subspecies as they are highly divergent both in sperm morphology and constitute a monophyletic clade in the mtDNA tree. Furthermore, this study shows that the *volgae* subspecies does not fall into one group, but that individuals from this (putative) subspecies are distributed across both mtDNA and nuclear gene trees. These findings, combined with the fact that the sperm measurements of *volgae* males fall between those of *svecica* and *cyanecula* males, and that males in our study population show a mixture of

ornamental spot colouration (chestnut, white and a chestnut/white mix; E. Matsyna, *unpublished data*), renders the status of *volgae* as a separate subspecies questionable.

Sperm divergence and speciation

Given the apparently rapid divergence of sperm morphology in bluethroats and some other taxa and the comparative evidence that sperm competition leads to both stabilizing and directional selection on sperm morphology, a plausible hypothesis is that sperm competition increases the likelihood of sperm divergence in allopatric populations, which in turn may result in reproductive isolation upon secondary contact and, ultimately, speciation (Howard et al., 2009; Pitnick et al., 2009). If so, one would predict that species with high sperm competition should show higher levels of sperm divergence than those with low sperm competition, and that speciation should be more rapid in species with high sperm competition than in those with low sperm competition. While the last prediction is difficult to evaluate at present due to lack of data in the literature, the first prediction is supported by available data on within-species sperm divergence in passerines, although the data points are few (table 4): the two species with small and non-significant sperm divergence both have low levels of sperm competition (pied flycatcher: Lifjeld et al. 1991, 15% of nests, 4% of nestlings; common redstart: 11% of nests, 2% of nestlings; Kleven *et al.* 2007) compared to the four species with significant sperm divergence (% nests, range: 48-72, % nestlings, range: 26-42) (Weatherhead and Boag 1995; Johnsen and Lifjeld 2003; Kleven et al. 2005; Schmoll et al. 2005,). The hypothesis that sperm competition promotes speciation through its effects on sperm divergence clearly warrants further investigation.

In conclusion, high divergence in sperm morphology combined with low genetic divergence suggests that sperm morphology has evolved rapidly in the bluethroat subspecies complex. While the relative importance of selection (e.g. via sperm competition) and drift is unknown, we suggest that selection is likely to have played an important role in driving sperm evolution because of the relatively short time span over which change has occurred. Moreover, our findings suggest that sperm divergence may play an important role in the early stages of the speciation process. Divergences in both primary (this study) and secondary sexual characters (song: Turcokova et al. 2010; Turcokova et al. 2011; plumage: Johnsen et al. 2006) among several distinct bluethroats subspecies (Johnsen et al. 2006), suggests a role for sexual selection in the diversification of this subspecies complex.

Acknowledgement

We thank Marc Naguib for providing us with blood samples of *Luscinia megarhynchos*, Patrick Bonnet, Sophie Questiau and Matthieu Marquet for help in the field, Eirik Rindal for statistical support, and Melissah Rowe, Becky Cramer and Gunnhild Marthinsen for helpful comments on the manuscript. Funding was provided by the Natural History Museum, University of Oslo (PhD fellowship to SH).

References

Andersson M (1994) *Sexual selection*. Princeton University Press, Princeton, New Jersey.

Balshine S, Leach BJ, Neat F, Werner NY and Montgomerie R (2001) Sperm size of African cichlids in relation to sperm competition. *Behavioral Ecology* 12: 726-731.

Birkhead TR and Møller AP (1998) *Sperm competition and sexual selection*. Academic Press, San Diego, Calif.

Birkhead TR, Pellatt EJ, Brekke P, Yeates R and Castillo-Juarez H (2005) Genetic effects on sperm design in the zebra finch. *Nature* 434: 383-387.

Borge ., Webster MT, Andersson G and Sætre GP (2005) Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* 171: 1861-1873.

Breed WG and Taylor J (2000) Body mass, testes mass, and sperm size in murine rodents. *Journal of Mammalogy* 81: 758-768.

Briskie JV, Montgomerie R and Birkhead TR (1997) The evolution of sperm size in birds. *Evolution* 51: 937-945.

Bromham L and Penny D (2003) The modern molecular clock. *Nature Reviews Genetics* 4: 216-224.

Byrne PG, Roberts JD and Simmons LW (2002) Sperm competition selects for increased testes mass in Australian frogs. *Journal of Evolutionary Biology* 15: 347-355.

Byrne PG, Simmons LW and Roberts JD (2003) Sperm competition and the evolution of gamete morphology in frogs. *Proceedings of the Royal Society B-Biological Sciences* 270: 2079-2086.

Calhim S, Immler S and Birkhead TR (2007) Postcopulatory Sexual Selection Is Associated with Reduced Variation in Sperm Morphology. *PLoS ONE* 2: e413.

Clements JF, Schulenberg TS, Iliff MJ, Sullivan BL, Wood CL and Roberson D (2012) The eBird/Clements checklist of birds of the world: Version 6.7. Downloaded from http://www.birds.cornell.edu/clementschecklist/downloadable-clements-checklist

Cohen J (1977) *Reproduction*. Butterworths, London.

Coyne JA and Orr AH (2004) *Speciation*. Sinauer Associates, Inc., Sunderland.

Cramer ERA, Laskemoen T, Kleven O and Lifjeld JT (2012) Sperm length variation in House Wrens *Troglodytes aedon*. *Journal of Ornithology* doi: 10.1007/s10336-012-0878-3

Cramp S (1988) Bluethroat. In: *The Birds of the Western Palearcitc. Handbook of Birds in Europe, the Middle East and Africa*, (Cramp, S., ed.). pp. 645-661. Oxford University Press, Oxford.

del Hoyo J, Elliot A and Christie DA (2005) *Handbook of the birds of the world*. Lynx Edicions, Barcelona.

Dixon A and Birkhead TR (1997) Reproductive anatomy of the Reed Bunting: A species which exhibits a high degree of sperm competition through extra-pair copulations. *Condor* 99: 966-969.

Dray S and Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22: 1-20.

Excoffier L, Laval G and Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47-50.

Gage MJG (1994) Associations between body-size, mating pettern, testis size and sperm lengths across butterflies. *Proceedings of the Royal Society B-Biological Sciences* 258: 247-254.

Gage MJG and Freckleton RP (2003) Relative testis size and sperm morphometry across mammals: no evidence for an association between sperm competition and sperm length. *Proceedings of the Royal Society B-Biological Sciences* 270: 625-632.

Garcia-Gonzalez F and Simmons LW (2007) Shorter sperm confer higher competitive fertilization success. *Evolution* 61: 816-824.

Gill F and Donsker D (Eds) (2012) IOC World Bird Names (v 3.2). Available at http://www.worldbirdnames.org [Accessed *29.11.2012*].

Gomendio M and Roldan ERS (1991) Sperm competition influences sperm size in mammals. *Proceedings of the Royal Society B-Biological Sciences* 243: 181-185.

Grether GF (2010) Sexual Selection and Speciation. Encyclopedia of Animal Behavior. D. B. Michael and M. Janice. Oxford, Academic Press**:** 177-183.

Griffith SC, Owens IPF and Thuman KA (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology* 11: 2195-2212.

Harcourt AH, Purvis A and Liles L (1995) Sperm competition: mating system, not breeding season, affects testes size of primates. *Functional Ecology* 9: 468-476.

Higginson DM, Miller KB, Segraves KA and Pitnick S (2012) Female reproductive tract form drives the evolution of complex sperm morphology. *Proceedings of the National Academy of Sciences of the United States of America* 109: 4538-4543.

Hogner S, Laskemoen T, Lifjeld JT, Porkert J, Kleven O, Albayrak T, Kabasakal B and Johnsen A (in press) Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. *Ecology and Evolution*: 10.1002/ece3.398.

Hosken DJ (1997) Sperm competition in bats. *Proceedings of the Royal Society B-Biological Sciences* 264: 385-392.

Howard DJ, Paulumbi SR, Birge LM and Manier MK (2009) Sperm and speciation. In: *Sperm biology: An evolutionary perspective*, (Birkhead TR, Hosken DJ and Pitnick S, eds.). pp. 367-403. Academic Press, Oxford.

Hudson RR, Kreitman M and Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153-159.

Immler S and Birkhead TR (2007) Sperm competition and sperm midpiece size: no consistent pattern in passerine birds. *Proceedings of the Royal Society B-Biological Sciences* 274: 561-568.

Johnsen A, Andersson S, Fernandez JG, Kempenaers B, Pavel V, Questiau S, Raess M Rindal E and Lifjeld JT (2006) Molecular and phenotypic divergence in the bluethroat (*Luscinia svecica*) subspecies complex. *Molecular Ecology* 15: 4033-4047.

Johnsen A and Lifjeld JT (1995) Unattractive Males Guard Their Mates More Closely: an Experiment with Bluethroats (Aves, Turdidae: Luscinia s. svecica). *Ethology* 101: 200-212.

Johnsen A and Lifjeld JT (2003) Ecological constraints on extra-pair paternity in the bluethroat. *Oecologia* 136: 476-483.

Johnsen A, Rindal E, Ericson P, Zuccon D, Kerr K, Stoeckle M and Lifjeld JT (2010) DNA barcoding of Scandinavian birds reveals divergent lineages in trans-Atlantic species. *Journal of Ornithology* 151: 565-578.

Kleven O, Jacobsen F, Robertson RJ and Lifjeld JT (2005) Extrapair mating between relatives in the barn swallow: a role for kin selection? *Biology Letters* 1: 389-392.

Kleven O, Øigarden T, Foyn BE, Moksnes A, Røskaft E, Rudolfsen G, Stokke BG and Lifjeld JT (2007) Low frequency of extrapair paternity in the common redstart (*Phoenicurus phoenicurus*). *Journal of Ornithology* 148: 373-378.

Kleven O, Laskemoen T, Fossøy F, Robertson RJ and Lifjeld JT (2008) Intraspecific variation in sperm length is negatively related to sperm competition in passerine birds. *Evolution* 62: 494-499.

Kleven O, Fossøy F, Laskemoen T, Robertson RJ, Rudolfsen G and Lifjeld JT (2009) Comparative evidence for the evolution of sperm swimming speed by sperm competition and female sperm storage duration in passerine birds. *Evolution* 63: 2466-2473.

Kleven O, Laskemoen T and Lifjeld JT (2009) Sperm length in sand martins *Riparia riparia*: a comment on Helfenstein et al. *Journal of Avian Biology* 40: 241-242.

Landry C, Geyer LB, Arakaki Y, Uehara T and Palumbi SR (2003) Recent speciation in the Indo-West Pacific: rapid evolution of gamete recognition and sperm morphology in cryptic species of sea urchin. *Proceedings of the Royal Society B-Biological Sciences* 270: 1839-1847.

Laskemoen T, Albrecht T, Bonisoli-Alquati A, Cepak, J., De Lope F, Hermosell I G *et al.* (in press) Variation in sperm morphometry and sperm competition among barn swallow (*Hirundo rustica*) populations. *Behavioral Ecology and Sociobiology*: DOI: 10.1007/s00265-012-1450-0.

Laskemoen T, Kleven O, Fossøy F and Lifjeld JT (2007) Intraspecific variation in sperm length in two passerine species, the bluethroat *Luscinia svecica* and the willow warbler *Phylloscopus trochilus*. . *Ornis Fennica* 84: 131-139.

Laskemoen T, Kleven O, Fossøy F, Robertson RJ, Rudolfsen G and Lifjeld JT (2010) Sperm quantity and quality effects on fertilization success in a highly promiscuous passerine, the tree swallow *Tachycineta bicolor*. *Behavioral Ecology and Sociobiology* 64: 1473-1483.

Leinonen T, Cano JM, Mäkinen H and Merilä J (2006) Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology* 19: 1803-1812.

Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

Lifjeld JT, Slagsvold T and Lampe HM (1991) Low-frequency of extra-pair paternity in pied flycatchers revealed by DNA fingerprinting. *Behavioral Ecology and Sociobiology* 29: 95-101.

Lifjeld JT, Hoenen A, Johannessen LE, Laskemoen T, Lopes R, Rodrigues P and Rowe M (2012a) The Azores bullfinch (*Pyrrhula murina*) has the same unusual and size-variable sperm morphology as the Eurasian bullfinch (*Phyrrhula phyrrula*). *Biological Journal of the Linnean Society* in press.

Lifjeld JT, Laskemoen T, Kleven O, Albrecht T and Robertson RJ (2010) Sperm Length Variation as a Predictor of Extrapair Paternity in Passerine Birds. *PLoS ONE* 5: e13456.

Lifjeld JT, Laskemoen T, Kleven O, Pedersen ATM, Lampe HM, Rudolfsen G, Schmoll T and Slagsvold T (2012b) No evidence for pre-copulatory sexual selection in a passerine bird. *PLoS ONE* 7.

Lohman DJ, Prawiradilaga DM and Meier R (2009) Improved COI barcoding primers for Southeast Asian perching birds (Aves: Passeriformes). *Molecular Ecology Resources* 9: 37-40.

Lüpold S, Calhim S, Immler S and Birkhead TR (2009) Sperm morphology and sperm velocity in passerine birds. *Proceedings of the Royal Society B-Biological Sciences* 276: 1175-1181.

Lüpold S, Linz G and Birkhead T (2009b) Sperm design and variation in the New World blackbirds (Icteridae). *Behavioral Ecology and Sociobiology* 63: 899-909.

Lüpold S, Westneat DF and Birkhead TR (2011) Geographical variation in sperm morphology in the red-winged blackbird (*Agelaius phoeniceus*). *Evolutionary Ecology* 25: 373-390.

Maan ME and Seehausen O (2011) Ecology, sexual selection and speciation. *Ecology Letters* 14: 591-602.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209-220.

Mayaud, N. (1958) La Gorge-Bleu à miroir (*Luscinia svecica*) en Europe. Evolution des populations. Zones d`hivernage. *Alauda* 26: 11.

Morrow EH and Gage MJG (2000) The evolution of sperm length in moths. *Proceedings of the Royal Society B-Biological Sciences* 267: 307-313.

Morrow EH and Gage MJG (2001) Artificial selection and heritability of sperm length in *Gryllus bimaculatus*. *Heredity* 87: 356-362.

Møller AP (1991) Sperm Competition, Sperm Depletion, Paternal Care, and Relative Testis Size in Birds. *The American Naturalist* 137: 882-906.

Panhuis TM, Butlin R, Zuk M.and Tregenza T (2001) Sexual selection and speciation. *Trends in Ecology & Evolution* 16: 364-371.

Peiponen V A (1960) Verhaltensstudien am Blaukehlchen (*Luscinia s. svecica*). *Ornis Fennica* 37: 69-83.

Pitnick S, Hosken DJ and Birkhead TR (2009) Sperm morphological diversity. In: *Sperm biology: an evolutionary perspective* (Birkhead TR, Hosken DJ and Pitnick S, eds.). pp. 69-149. Academic Press, Oxford.

Pizzari, T. & Parker, G. A. (2009) Sperm competition and sperm phenotype. In: *Sperm Biology: an evolutionary perspective* (Birkhead, T. R., Hosken, D. J. & Pitnick, S., eds.). pp. 206-245. Academic Press, Oxford.

Pulquério MJF and Nichols RA (2007) Dates from the molecular clock: how wrong can we be? *Trends in Ecology & Evolution* 22: 180-184.

Päckert M, Martens J, Tietze DT, Dietzen C, Wink M and Kvist L (2007) Calibration of a molecular clock in tits (Paridae) - Do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? *Molecular Phylogenetics and Evolution* 44: 1-14.

Questiau S, Eybert MC, Gaginskaya AR, Gielly L and Taberlet P (1998) Recent divergence between two morphologically differentiated subspecies of bluethroat (Aves : Muscicapidae : *Luscinia svecica*) inferred from mitochondrial DNA sequence variation. *Molecular Ecology* 7: 239-245.

Questiau S, Eybert MC and Taberlet P (1999) Amplified fragment length polymorphism (AFLP) markers reveal extra-pair parentage in a bird species: the bluethroat (*Luscinia svecica*). *Molecular Ecology* 8: 1331-1339.

R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rowe M and Pruett-Jones S (2011) Sperm competition selects for sperm quantity and quality in the Australian Maluridae. *PLoS ONE* 6.

Schmoll T, Dietrich V, Winkel W, Epplen JT, Schurr F and Lubjuhn T (2005) Paternal genetic effects on offspring fitness are context dependent within the extrapair mating system of a socially monogamous passerine. *Evolution* 59: 645-657.

Schmoll T and Kleven O (2011) Sperm dimensions differ between two Coal Tit Periparus ater populations. *Journal of Ornithology* 152: 515-520.

Sokal RR and Rohlf FJ (1995) *The Principles and Practice of Statistics in Biological Research,* 3 ed. W H Freeman & Co, New York.

Stockley P, Gage MJG, Parker GA and Møller AP (1997) Sperm competition in fishes: the evolution of testis size and ejaculate characteristics. *American Naturalist* 149: 933-954.

Tajima F (1989) Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28: 2731-2739.

Tarr CL (1995) Primers for amplification and determination of mitochondrial control-region sequences in oscine passerines. *Molecular Ecology* 4: 527-529.

Turcokova L, Osiejuk TS, Pavel V, Glapan J and Petruskova T (2010) Song divergence of two Bluethroat subspecies (*Luscinia s. svecica* and *cyanecula*). *Ornis Fennica* 87: 168-179.

Turcokova L, Pavel V, Chutny B, Petrusek A and Petruskova T (2011) Differential response of males of a subarctic population of Bluethroat *Luscinina svecica svecica* to playbacks of their own and foreign subspecies. *Journal of Ornithology* 152: 975-982.

Weatherhead PJ and Boag PT (1995) Pair and extra-pair mating success relative to male quality in red-winged blackbirds. *Behavioral Ecology and Sociobiology* 37: 81-91.

Weir BS and Cockerham CC (1984) Estimating f-statistics for the analysis of population-structure. *Evolution* 38: 1358-1370.

Weir JT and Schluter D (2008) Calibrating the avian molecular clock. *Molecular Ecology* 17: 2321-2328.

Wolfson A (1952) The cloacal protuberance: a means for determining breeding condition in live male passerines. *Bird Banding* 23: 159-165.

Woolley DM (1970) Selection for the length of the spermatozoan midpiece in the mouse. *Genetics Research* 16: 261-275.

Zink RM, Drovetski SV, Questiau S, Fadeev IV, Nesterov EV, Westberg MC *et al.* (2003) Recent evolutionary history of the bluethroat (*Luscinia svecica*) across Eurasia. *Molecular Ecology* 12: 3069-3075.

Table 1: Summary of the different sperm trait measures in the respective populations

	Head length		**Midpiece length**		**Tail length**		**Total length**		**CV_{bm}**	**CV_{wm}[a]**
	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range		
azuricollis (n=16)	19.28 ± 0.43	18.61-20.18	190.05 ± 7.88	171.62-205.18	15.88 ± 3.98	10.71-26.76	225.21 ± 7.40	213.76-239.79	3.30	1.31
cyanecula (n=14)	18.80 ± 0.59	18.06-19.86	167.79 ± 4.22	162.85-173.27	14.06 ± 2.34	10.24-18.37	200.64 ± 4.05	192.51-205.90	2.04	1.92
namnetum (n=14)	18.63± 0.70	17.45-19.72	173.81 ± 5.91	167.91-181.48	14.07 ±2.42	10.69-18.56	206.51 ± 5.92	198.05-214.62	2.89	1.52
svecica (n=29)	19.69 ± 0.75	18.85-21.07	175.96 ± 6.14	157.78-181.96	15.15 ± 3.59	13.06-19.86	210.79 ± 4.00	202.00-221.04	1.91	1.30
volgae (n=9)	19.46 ± 0.95	17.91-21.29	171.67 ± 7.32	161.39-179-38	14.87 ±4.21	7.45 -22.69	205.99 ± 5.00	197.70-213.55	2.45	1.52

[a] = mean CV within males

Table 2: Tukey HSD tests; variables: head, midpiece, total sperm length and CV_{wm}. Approximate Probabilities for Post Hoc Tests Error: Between MS = 27.80, df = 77.00. Significant results highlighted in bold (after Bonferroni corrections).

Trait	subspecies	*azuricollis*	*cyanecula*	*namnetum*	*svecica*
Head	*cyanecula*	0.33			
	namnetum	0.09	1.00		
	svecica	0.32	**0.002**	**<0.001**	
	volgae	0.97	0.18	0.05	0.90
Midpiece	*cyanecula*	**<0.001**			
	namnetum	**<0.001**	**0.002**		
	svecica	**<0.001**	0.40	0.61	
	volgae	**<0.001**	0.83	0.10	0.93
Total	*cyanecula*	**<0.001**			
	namnetum	**<0.001**	0.03		
	svecica	**<0.001**	**<0.001**	0.10	
	volgae	**<0.001**	0.13	1.00	0.13
CV_{wm}	*cyanecula*	**<0.001**			
	namnetum	0.48	0.03		
	svecica	1.00	**<0.001**	0.31	
	volgae	0.60	0.07	1.00	0.46

Table 3: F_{ST}-values based on mitochondrial DNA with associated P-values above the diagonal. Significant results highlighted in bold (after Bonferroni corrections).

	azuricollis	*cyanecula*	*namnetum*	*svecica*	*volgae*
azuricollis		**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**
cyanecula	0.80		**< 0.001**	**< 0.001**	**< 0.001**
namnetum	0.75	0.20		**< 0.001**	**< 0.001**
svecica	0.76	0.56	0.53		**< 0.001**
volgae	0.70	0.36	0.34	0.26	

Table 4: Intraspecific studies (between populations/subspecies).

Common name	Scientific name	Min	Max	Number of populations	Significant differentiation in sperm?	Relative difference(%)[a]	Source
Pied flycatcher	*Ficedula hypoleuca*	102.9	104.2	3	No	1.3	Lifjeld *et al.* (2012b)
Red-winged blackbird	*Agelaius phoeniceus*	140.0	146.0	17	Yes	2.9	Lüpold *et al.* (2011)
Barn swallow	*Hirundo rustica*	87.9	91.2	7	Yes	3.7	Laskemoen *et al.* (in press)
Coal tit	*Periparus ater*	91.7	94.6	2	Yes	3.1	Schmoll and Kleven (2011)
Common redstart	*Phoenicurus phoenicurus*	163.6	164.1	4	No	0.3	Hogner *et al.* (in press)
Bluethroat	*Luscinia svecica*	200.6	225.2	5	Yes	11.6	This study

[a] = Standardised difference in sperm length ((max-min)/mean), see methods.

Table 5. Comparison of differentiation in total sperm length between 4 species pairs of passerines.

Common name	Scientific name	Total length (µm)	SD (µm)	N	Difference (µm)[a]	Relative difference (%)[b]	Source	COI divergence (%) [c]
Black redstart	*Phoenicurus ochruros*	180.8	3.5	5			Laskemoen, Albrecht, Lifjeld unpublished data	
Common redstart	*Phoenicurus phoenicurus*	163.7	5.8	75	17.1	9.9	Hogner *et al.* (in press)	7.5
Collared flycatcher	*Ficedula albicollis*	101.2	2.4	14			Laskemoen, Albrecht, Lifjeld unpublished data	
Pied flycatcher	*Ficedula hypoleuca*	103.5	2.8	80	2.3	2.2	Lifjeld *et al.* (2012b)	2.1
Azores bullfinch	*Pyrrhula murina*	45.6	4.3	11			Lifjeld *et al.* (2012a)	
European bullfinch	*Pyrrhula pyrrhula*	46.3	4.3	13	0.7	1.5	Lifjeld *et al.* (2012a)	1.0
House sparrow	*Passer domesticus*	99.7	3.2	27			Laskemoen, Sætre, Johnsen, Lifjeld unpublished data	
Spanish sparrow	*Passer hispaniolensis*	100.0	2.4	16	0.3	0.3	Laskemoen, Sætre, Johnsen, Lifjeld unpublished data	2.8

[a] = Absolute difference in mean total sperm length

[b] = Standardised difference in sperm length ((max-min)/mean), see methods.

[c] = Johnsen, Lifjeld, Hogner unpublished data

Figure 1: Box plot illustrating differences in total sperm length for the five subspecies of the bluethroat. White line represents mean values, boxes indicate ± SE and whiskers indicate ± SD. Se results for test statistics.

Figure 2: Maximum likelihood tree (HKY+G model, 10,000 bootstrap replicates) based on the COI and control region combined (1229bp) for 84 bluethroats. Only bootstrap values above 50% are shown. Green = *azuricollis*, blue = *cyanecula*, red = *namnetum*, black = *svecica*, pink = *volgae* and purple = the outgroup *Luscinia megarhynchos*

Figure 3: Plot illustrating pair-wise correlation between the genetic variation (F_{ST}) in mtDNA and total sperm length variation (P_{ST}). See results for test statistics.

Figure 1

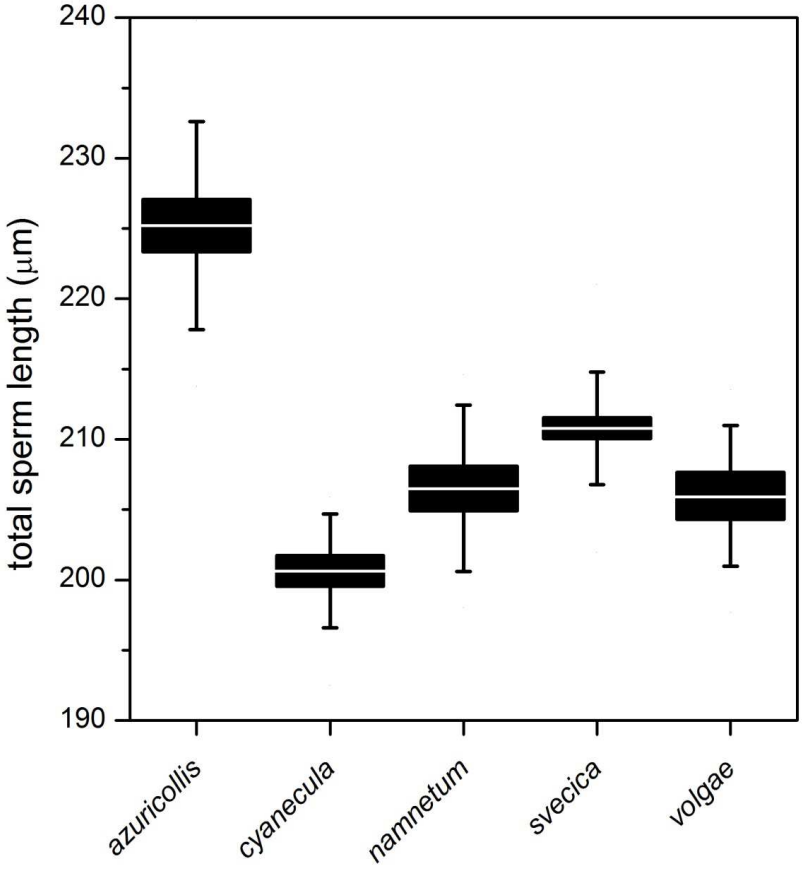

240
230
220
210
200
190
total sperm length (μm)
azuricollis
cyanecula
namnetum
svecica
volgae

Figure 2

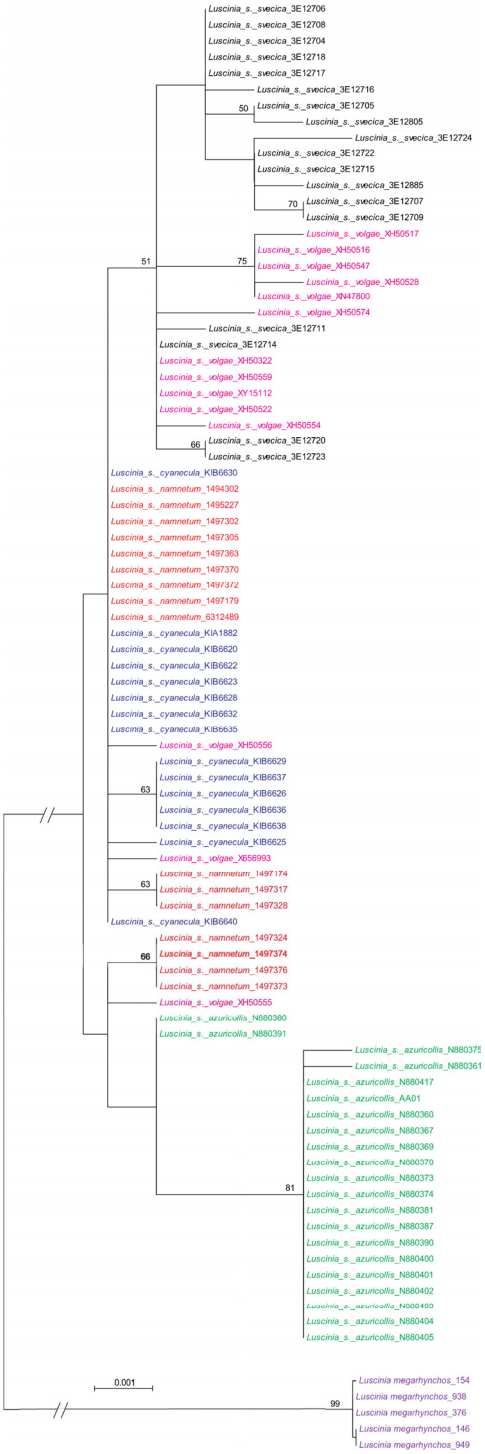

Luscinia_s._svecica_3E12706
Luscinia_s._svecica_3E12708
Luscinia_s._svecica_3E12704
Luscinia_s._svecica_3E12718
Luscinia_s._svecica_3E12717
Luscinia_s._svecica_3E12716
50
Luscinia_s._svecica_3E12705
Luscinia_s._svecica_3E12805
Luscinia_s._svecica_3E12724
Luscinia_s._svecica_3E12722
Luscinia_s._svecica_3E12715
Luscinia_s._svecica_3E12885
70
Luscinia_s._svecica_3E12707
Luscinia_s._svecica_3E12709
Luscinia_s._volgae_XH50517
Luscinia_s._volgae_XH50516
51
75
Luscinia_s._volgae_XH50547
Luscinia_s._volgae_XH50528
Luscinia_s._volgae_XN47800
Luscinia_s._volgae_XH50574
Luscinia_s._svecica_3E12711
Luscinia_s._svecica_3E12714
Luscinia_s._volgae_XH50322
Luscinia_s._volgae_XH50559
Luscinia_s._volgae_XY15112
Luscinia_s._volgae_XH50522
Luscinia_s._volgae_XH50554
66
Luscinia_s._svecica_3E12720
Luscinia_s._svecica_3E12723
Luscinia_s._cyanecula_KIB6630
Luscinia_s._namnetum_1494302
Luscinia_s._namnetum_1495227
Luscinia_s._namnetum_1497302
Luscinia_s._namnetum_1497305
Luscinia_s._namnetum_1497363
Luscinia_s._namnetum_1497370
Luscinia_s._namnetum_1497372
Luscinia_s._namnetum_1497179
Luscinia_s._namnetum_6312489
Luscinia_s._cyanecula_KIA1882
Luscinia_s._cyanecula_KIB6620
Luscinia_s._cyanecula_KIB6622
Luscinia_s._cyanecula_KIB6623
Luscinia_s._cyanecula_KIB6628
Luscinia_s._cyanecula_KIB6632
Luscinia_s._cyanecula_KIB6635
Luscinia_s._volgae_XH50556
Luscinia_s._cyanecula_KIB6629
Luscinia_s._cyanecula_KIB6637
63
Luscinia_s._cyanecula_KIB6626
Luscinia_s._cyanecula_KIB6636
Luscinia_s._cyanecula_KIB6638
Luscinia_s._cyanecula_KIB6625
Luscinia_s._volgae_X656993
Luscinia_s._namnetum_1497174
63
Luscinia_s._namnetum_1497317
Luscinia_s._namnetum_1497328
Luscinia_s._cyanecula_KIB6640
Luscinia_s._namnetum_1497324
66
Luscinia_s._namnetum_1497374
Luscinia_s._namnetum_1497376
Luscinia_s._namnetum_1497373
Luscinia_s._volgae_XH50555
Luscinia_s._azuricollis_N880380
Luscinia_s._azuricollis_N880391
Luscinia_s._azuricollis_N880375
Luscinia_s._azuricollis_N880361
Luscinia_s._azuricollis_N880417
Luscinia_s._azuricollis_AA01
Luscinia_s._azuricollis_N880360
Luscinia_s._azuricollis_N880367
Luscinia_s._azuricollis_N880369
Luscinia_s._azuricollis_N880370
Luscinia_s._azuricollis_N880373
81
Luscinia_s._azuricollis_N880374
Luscinia_s._azuricollis_N880381
Luscinia_s._azuricollis_N880387
Luscinia_s._azuricollis_N880390
Luscinia_s._azuricollis_N880400
Luscinia_s._azuricollis_N880401
Luscinia_s._azuricollis_N880402
Luscinia_s._azuricollis_N880403
Luscinia_s._azuricollis_N880404
Luscinia_s._azuricollis_N880405
0.001
Luscinia megarhynchos_154
99
Luscinia megarhynchos_938
Luscinia megarhynchos_376
Luscinia megarhynchos_146
Luscinia megarhynchos_949

Figure 3

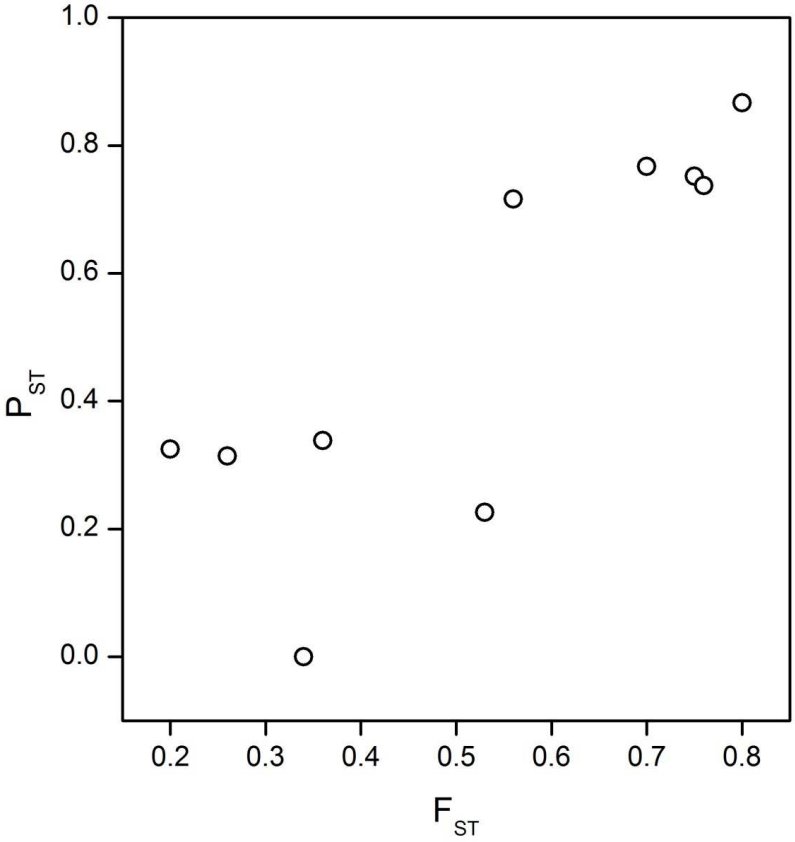


Table S1: Overview of the populations sampled and their associated coordinates.

Subspecies	Location	Latitude	Longitude
azuricollis	León, Spain	42 º22'N	5º59'E
cyanecula	Wroclaw, Poland	51º11'N	16º57'E
cyanecula	Thüringen, Germany	50º21'N	10º44'E
cyanecula	Trebon, Czech R.	49º3'N	14º43'E
namnetum	Brière, France	47º21'N	2º12'W
namnetum	Guérande, France	47º20'N	2º25'W
svecica	Heimdalen, Norway	61º25'N	8º52'E
volgae	Nizhny Novgorod, Russia	56º15'N	44º9'E

Table S2: Overview of the individuals sampled, and their inclusion in sperm, nuclear or mitochondrial analyses.

J. No	Band No	Scientific subspecies	Sex	Age	Date	Locality, Country	Position LatLon	Sperm	mtDNA	Intron
1214	9707599	*Luscinia s. svecica*	Female	2K	25.05.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1152	E141453	*Luscinia s. svecica*	Female	2K	27.05.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1197	E141473	*Luscinia s. svecica*	Female	2K	20.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1228	E141475	*Luscinia s. svecica*	Female	3K+	22.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1165	9707643	*Luscinia s. svecica*	Female	n.a.	23.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1273	E141479	*Luscinia s.svecica*	Female	3K+	24.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1179	E141474	*Luscinia s. svecica*	Female	3K+	22.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1188	E08002	*Luscinia s. svecica*	Female	2K	25.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1236	E08008	*Luscinia s. svecica*	Female	2K	25.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1250	E141486	*Luscinia s. svecica*	Female	2K	26.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1264	E08021	*Luscinia s. svecica*	Female	3K+	26.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1256	E08038	*Luscinia s. svecica*	Female	3K+	29.06.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1103	E08162	*Luscinia s. svecica*	Female	2K	01.07.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1281	E141498	*Luscinia s. svecica*	Female	2K	01.07.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
1242	E08305	*Luscinia s. svecica*	Female	3K+	05.07.1998	Øvre Heimdalen, Norway	61°25'N, 8°52'E			X
11065	S109821	*Luscinia s. cyanecula*	Female	3K+	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11070	T838187	*Luscinia s. cyanecula*	Female	3K+	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11072	T839101	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11073	T839102	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11074	T839103	*Luscinia s. cyanecula*	Female	3K+	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11077	T986965	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11081	T991856	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11082	T991857	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11084	T991859	*Luscinia s. cyanecula*	Female	3K+	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11086	T991861	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11087	T991862	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11094	T992851	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X
11096	n.a	*Luscinia s. cyanecula*	Female	2K	04.04.2000	Trebonsko, Czech Republic	49°3'N, 14°43'E			X

11223	L381711	*Luscinia s. azuricollis*	Female	3K+	02.06.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11224	L381720	*Luscinia s. azuricollis*	Female	3K+	30.06.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11225	L381725	*Luscinia s. azuricollis*	Female	2K	13.07.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11226	L381726	*Luscinia s. azuricollis*	Female	3K+	13.07.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11227	L381723	*Luscinia s. azuricollis*	Female	1K	13.07.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11228	L381445	*Luscinia s. azuricollis*	Female	1K	30.07.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11229	L381446	*Luscinia s. azuricollis*	Female	1K	30.07.2000	Valduerna, Spain	42°20'N, 5°58'W			X
11319	939753	*Luscinia s. namnetum*	Female	Adult	27.05.1998	Guérande, France	47°20'N, 2°25'W			X
11322	1152242	*Luscinia s.namnetum*	Female	Adult	08.04.1998	Guérande, France	47°20'N, 2°25'W			X
11326	1152463	*Luscinia s. namnetum*	Female	1K	23.05.1998	Guérande, France	47°20'N, 2°25'W			X
11329	1152473	*Luscinia s. namnetum*	Female	Adult	27.05.1998	Guérande, France	47°20'N, 2°25'W			X
11330	1219330	*Luscinia s. namnetum*	Female	1K	30.03.1998	Guérande, France	47°20'N, 2°25'W			X
11332	1219339	*Luscinia s. namnetum*	Female	Adult	08.04.1998	Guérande, France	47°20'N, 2°25'W			X
11334	1219385	*Luscinia s. namnetum*	Female	Adult	10.05.1998	Guérande, France	47°20'N, 2°25'W			X
11337	1219399	*Luscinia s. namnetum*	Female	1K	19.05.1998	Guérande, France	47°20'N, 2°25'W			X
11338	1219430	*Luscinia s. namnetum*	Female	1K	21.06.1998	Guérande, France	47°20'N, 2°25'W			X
11339	2551731	*Luscinia s. namnetum*	Female	Adult	07.05.1998	Guérande, France	47°20'N, 2°25'W			X
11346	VA99615	*Luscinia s. cyanecula*	Female	n.a	10.04.1996	Thüringen, Germany	50°21'N, 10°44'E			X
11349	VA99617	*Luscinia s. cyanecula*	Female	n.a	11.04.1996	Thüringen, Germany	50°21'N, 10°44'E			X
11350	BY00081	*Luscinia s. cyanecula*	Female	n.a	13.04.1996	Thüringen, Germany	50°9'N, 11°10'E			X
11352	VA99618	*Luscinia s. cyanecula*	Female	n.a	15.04.1996	Thüringen, Germany	50°21'N, 10°44'E			X
11355	BE70614	*Luscinia s. cyanecula*	Female	n.a	19.04.1996	Thüringen, Germany	50°9'N, 11°10'E			X
11356	BE70618	*Luscinia s. cyanecula*	Female	n.a	19.04.1996	Thüringen, Germany	50°9'N, 11°10'E			X
11357	VA99621	*Luscinia s.cyanecula*	Female	n.a	20.04.1996	Thüringen, Germany	50°21'N, 10°44'E			X
11460	L677555	*Luscinia s. azuricollis*	Male	3K+	28.06.2003	Palacios, Spain	42°20'N, 5°58'W			X
11469	L833335	*Luscinia s. azuricollis*	Female	1K	20.09.2003	Valporquero, Spain	42°20'N, 5°58'W			X
22448	3E11746	*Luscinia s. svecica*	Male	3K+	06.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X		
22450	3E11748	*Luscinia s. svecica*	Male	n.a	07.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X		
22461	3E11759	*Luscinia s. svecica*	Male	n.a	08.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X		
22471	3E11769	*Luscinia s. svecica*	Male	n.a	09.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X		
22472	3E11770	*Luscinia s. svecica*	Male	n.a	09.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X		

22475	3E11773	*Luscinia s. svecica*	Male	n.a	10.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
22479	3E11777	*Luscinia s. svecica*	Male	n.a	11.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
22481	3E11779	*Luscinia s. svecica*	Male	n.a	11.06.2007	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
25709	3E11575	*Luscinia s. svecica*	Male	2k	04.06.2008	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
32291	3E12805	*Luscinia s. svecica*	Female	3K+	26.06.2008	Øvre Heimdalen, Norway	61°25'N, 8°53'E		X
29102	3E12239	*Luscinia s. svecica*	Male	2k+	05.06.2009	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
29107	3E12244	*Luscinia s. svecica*	Male	2k+	06.06.2009	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
29109	3E12247	*Luscinia s. svecica*	Male	2k+	06.06.2009	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
29116	3E12253	*Luscinia s. svecica*	Male	2k+	06.06.2009	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
835*	3E12263	*Luscinia s. svecica*	Male	2k+	07.06.2009	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
29122	3E12260	*Luscinia s. svecica*	Male	2k+	07.06.2009	Øvre Heimdalen, Norway	61°25'N, 8°53'E	X	
32484	3E11572	*Luscinia s. svecica*	Male	n.a	04.06.2008	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
34409	1497302	*Luscinia s. namnetum*	Male	2K	03.04.2011	Briere, France	47°21'3N,2°12'W	X	X
34415	1497317	*Luscinia s. namnetum*	Male	2K	03.04.2011	Briere, France	47°21'N, 2°12'W		X
34416	6312489	*Luscinia s. namnetum*	Male	2K	03.04.2011	Briere, France	47°21'N, 2°12'W		X
34417	1495227	*Luscinia s. namnetum*	Male	3K+	03.04.2011	Briere, France	47°21'N, 2°12'W		X
34420	1497324	*Luscinia s. namnetum*	Male	3K+	03.04.2011	Briere, France	47°21'N, 2°12'W	X	X
34422	1497328	*Luscinia s. namnetum*	Male	3K+	03.04.2011	Briere, France	47°21'N, 2°12'W	X	X
34423	1497323	*Luscinia s. namnetum*	Male	3K+	03.04.2011	Briere, France	47°21'N, 2°12'W	x	
34426	1497305	*Luscinia s. namnetum*	Male	3K+	03.04.2011	Briere, France	47°21'N, 2°12'W	X	X
34427	1497174	*Luscinia s. namnetum*	Male	2K	03.04.2011	Briere, France	47°21'N, 2°12'W		X
34428	1497179	*Luscinia s. namnetum*	Male	3K+	03.04.2011	Briere, France	47°21'N, 2°12'W		X
34436	1497356	*Luscinia s. namnetum*	Male	2K	04.04.2011	Briere, France	47°21'N, 2°12'W	X	
34437	1497363	*Luscinia s. namnetum*	Male	3K+	04.04.2011	Briere, France	47°21'N, 2°12'W		X
40805	1497380	*Luscinia s. namnetum*	Male	2K	04.05.2011	Guerande, France	47°17'N, 2°28'W	X	
40808	1497398	*Luscinia s. namnetum*	Male	3K+	07.06.2011	Guerande, France	47°17'N, 2°28'W	X	
34445	1497370	*Luscinia s. namnetum*	Male	3K+	05.04.2011	Guerande, France	47°17'N, 2°28'W	X	X
34446	1497372	*Luscinia s. namnetum*	Male	2K	05.04.2011	Guerande, France	47°17'N, 2°27'W		X
34447	1497373	*Luscinia s. namnetum*	Male	3K+	05.04.2011	Guerande, France	47°17'N, 2°27'W	X	X
34448	1497374	*Luscinia s. namnetum*	Male	3K+	05.04.2011	Guerande, France	47°17'N, 2°28'W	X	X
34449	1494302	*Luscinia s. namnetum*	Male	6K+	06.04.2011	Guerande, France	47°17'N, 2°28'W	X	X

34451	1497376	*Luscinia s. namnetum*	Male	n.a	06.04.2011	Guerande, France	47°17'N, 2°28'W	X	X
34452	1497181	*Luscinia s. namnetum*	Male	n.a	03.04.2011	Briere, France	47°21'N, 2°12'W	X	
37429	3E12704	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E		X
37430	3E12705	*Luscinia s. svecica*	Male	2K	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37431	3E12706	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37432	3E12707	*Luscinia s. svecica*	Male	2K	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E		X
37433	3E12708	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37434	3E12709	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37436	3E12711	*Luscinia s. svecica*	Male	2K	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37437	3E12712	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
37439	3E12714	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37440	3E12715	*Luscinia s. svecica*	Male	2K	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37441	3E12716	*Luscinia s. svecica*	Male	3K+	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E		X
37442	3E12717	*Luscinia s. svecica*	Male	2K	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E		X
37443	3E12718	*Luscinia s. svecica*	Male	3K+	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37445	3E12720	*Luscinia s. svecica*	Male	3K+	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	X
37446	3E12721	*Luscinia s. svecica*	Male	3K+	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
37447	3E12722	*Luscinia s. svecica*	Male	2K	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E		X
37448	3E12723	*Luscinia s. svecica*	Male	n.a	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E		X
37449	3E12724	*Luscinia s. svecica*	Male	n.a	27.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
37451	3E12885	*Luscinia s. svecica*	Male	3K+	26.05.2010	Øvre Heimdalen, Norway	61°25'N, 8°52'E	X	
38659	K1B6620	*Luscinia s. cyanecula*	Male	3K+	16.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38661	K1B6622	*Luscinia s. cyanecula*	Male	3K+	16.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38662	K1B6623	*Luscinia s. cyanecula*	Male	2K	17.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38665	K1B6626	*Luscinia s. cyanecula*	Male	3K+	17.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38667	K1B6628	*Luscinia s. cyanecula*	Male	3K+	17.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38668	K1B6629	*Luscinia s. cyanecula*	Male	2K	17.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38669	K1B6630	*Luscinia s. cyanecula*	Male	3K+	18.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38671	K1A1882	*Luscinia s. cyanecula*	Male	3K+	18.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38672	K1B6632	*Luscinia s. cyanecula*	Male	3K+	18.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38675	K1B6635	*Luscinia s. cyanecula*	Male	2K	18.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X

38676	K1B6636	*Luscinia s. cyanecula*	Male	3K+	18.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38677	K1B6637	*Luscinia s. cyanecula*	Male	3K+	18.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38679	K1B6638	*Luscinia s. cyanecula*	Male	2K	19.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
38681	K1B6640	*Luscinia s. cyanecula*	Male	3K+	19.04.2009	Wroclaw, Poland	51°11'N, 16°57'E		X
6582*	K1E6027	*Luscinia s. cyanecula*	Male	3K+	24.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6583*	K1B6633	*Luscinia s. cyanecula*	Male	3K+	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6584*	K1B6635	*Luscinia s. cyanecula*	Male	3K+	24.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6585*	K1B6643	*Luscinia s. cyanecula*	Male	3K+	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6586*	K1E6011	*Luscinia s. cyanecula*	Male	3K+	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6587*	K1E6012	*Luscinia s. cyanecula*	Male	2K	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6588*	K1E6014	*Luscinia s. cyanecula*	Male	3K+	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6589*	K1E6017	*Luscinia s. cyanecula*	Male	3K+	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6590*	K1E6021	*Luscinia s. cyanecula*	Male	3K+	23.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6591*	K1E6023	*Luscinia s. cyanecula*	Male	2K	24.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6592*	K1E6025	*Luscinia s. cyanecula*	Male	3K+	24.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6593*	K1E6028	*Luscinia s. cyanecula*	Male	3K+	24.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6594*	K1E6031	*Luscinia s. cyanecula*	Male	2K	25.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6595*	K1E6032	*Luscinia s. cyanecula*	Male	2K	25.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
6596*	K4B4313	*Luscinia s. cyanecula*	Male	3K+	24.04.2010	Wroclaw, Poland	51°11'N, 16°57'E	X	
38683	N880360	*Luscinia s. azuricollis*	Male	3K+	03.06.2010	Rodillazo, Spain	42°54'N, 5°28'W		X
38684	N880361	*Luscinia s. azuricollis*	Male	2K	03.06.2010	Rodillazo, Spain	42°54'N, 5°28'W		X
38685	N880367	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38686	N880368	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38687	N880369	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38688	N880370	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38689	N880373	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38690	N880374	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38691	N880375	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W		X
38692	AA01-	*Luscinia s. azuricollis*	Male	3K+	12.06.2010	Bustos, Spain	42°22'N, 6°0'W	X	X
38693	N880380	*Luscinia s. azuricollis*	Male	3K+	14.06.2010	Pobladura de la Sierra, Spain	42°25'N, 6°29'W		X
38694	N880381	*Luscinia s. azuricollis*	Male	2K	14.06.2010	Pobladura de la Sierra, Spain	42°25'N, 6°29'W		X

38695	N880387	*Luscinia s. azuricollis*	Male	3K+	23.06.2010	Corcos, Spain	42º40'N, 5º5'W	X	X
38696	N880390	*Luscinia s. azuricollis*	Male	2K	23.06.2010	Corcos, Spain	42º40'N, 5º5'W		X
38697	N880391	*Luscinia s. azuricollis*	Male	3K+	23.06.2010	Corcos, Spain	42º40'N, 5º5'W	X	x
38698	N880399	*Luscinia s. azuricollis*	Male	2K	26.06.2010	Villar de Golfer, Spain	42º18'N, 6º6'W	X	
38699	N880400	*Luscinia s. azuricollis*	Male	3K+	26.06.2010	Villar de Golfer, Spain	42º19'N, 5º6'W	X	X
38700	N880401	*Luscinia s. azuricollis*	Male	3K+	26.06.2010	Villar de Golfer, Spain	42º19'N, 6º6'W	X	X
38701	N880402	*Luscinia s. azuricollis*	Male	2K	29.06.2010	Ferreras de Cepeda, Spain	42º37'N, 5º5'W	X	X
38702	N880403	*Luscinia s. azuricollis*	Male	3K+	29.06.2010	Ferreras de Cepeda, Spain	42º37'N, 5º5'W	X	X
38703	N880404	*Luscinia s. azuricollis*	Male	2K	29.06.2010	Ferreras de Cepeda, Spain	42º37'N, 5º5'W		X
38704	N880405	*Luscinia s. azuricollis*	Male	2K	29.06.2010	Ferreras de Cepeda, Spain	42º37'N, 5º5'W	X	X
38705	N880417	*Luscinia s. azuricollis*	Male	3K+	29.06.2010	Ferreras de Cepeda, Spain	42º37'N, 5º5'W	X	X
40743	XH50516	*Luscinia s. volgae*	Male	2K	29.05.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E		X
40744	XH50517	*Luscinia s. volgae*	Male	3K+	29.05.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E		X
40745	XH50522	*Luscinia s. volgae*	Male	3K+	29.05.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40746	XH50528	*Luscinia s. volgae*	Male	3K+	06.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40747	XH50547	*Luscinia s. volgae*	Male	2K	13.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	x
40748	XH50554	*Luscinia s. volgae*	Male	3K+	18.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E		X
40749	XH50555	*Luscinia s. volgae*	Male	3K+	19.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40750	XH50556	*Luscinia s. volgae*	Male	3K+	19.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40751	XH50559	*Luscinia s. volgae*	Male	3K+	19.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40752	XH50574	*Luscinia s. volgae*	Male	3K+	27.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40753	XG56993	*Luscinia s. volgae*	Male	2K	13.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X
40754	XN50322	*Luscinia s. volgae*	Male	2K	13.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E		X
40755	XN47800	*Luscinia s. volgae*	Male	2K	13.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E		X
40756	XY15112	*Luscinia s. volgae*	Male	3K+	19.06.2011	Nizhny Novgorod region, Russia	56º15'N, 44º9'E	X	X

:

Table S3: Overview of the primers used and their PCR annealing temperatures.

Locus	C_1	Primer sequence (5'-3')	PCR_2	Reference
Vldlr-7	Z	F: AGACCATGATCTCCAGCGCT	56	Borge *et al.* (2005)
		R: CCTTCCAGGTAGACATGATG		
Brm-15	Z	F: AGCACCTTTGAACAGTGGTT	56	Borge *et al.* (2005)
		R: TACTTTATGGAGACGACGGA		
COI Passer	m	F: CCAACCACAAAGACATCGGAACC	52	Lohman *et al* (2009)
COI Passer		R: GTAAACTTCTGGGTGACCAAAGAATC		
L437	m	F:CTCACGAGAACCGAGCTACT	52	Tarr (1995)
H1248		R: CATCTTCAGTGTCATGCT		

[1] DNA class: Z = Z-linked, m = mtDNA

[2] Annealing temperature

Table S4: The genetic diversity for the different subspecies and regions.

Gene	Location	Subspecies	N	L	π	Tajima's D
COI+ control region	mtDNA	*azuricollis*	21	1239	0.0006	-1.42
		cyanecula	15	1239	0.0005	-0.02
		namnetum	16	1239	0.0009	0.71
		svecica	18	1239	0.0016	-0.84
		volgae	14	1219	0.0022	-1.20
		combined	84	1219	0.0023	-1.31
BRM+VLDLR	z-region	*azuricollis*	8	932	0.0006	1.17
		cyanecula	20	923	0.0074	0.03
		namnetum	10	931	0.0038	1.11
		svecica	15	926	0.0063	-1.21
		combined	53	923	0.0061	-1.16

Figure S1: Maximum likelihood tree (Tamura-3-Parameter model, 10,000 bootstrap replicates) based on the intron BRM-15 (353bp) for 53 bluethroats. Only bootstrap values above 50% are shown. Green = *azuricollis*, blue = *cyanecula*, red = *namnetum*, black = *svecica*, and purple = the outgroup *Luscinia megarhynchos*

Figure S2: Maximum likelihood tree (Kimura-2-Parameter model, 10,000 bootstrap replicates) based on the intron VLDLR (580bp) for 53 bluethroats. Only bootstrap values above 50% are shown. Green = *azuricollis*, blue = *cyanecula*, red = *namnetum*, black = *svecica*, and purple = the outgroup *Luscinia megarhynchos*

Figure S1

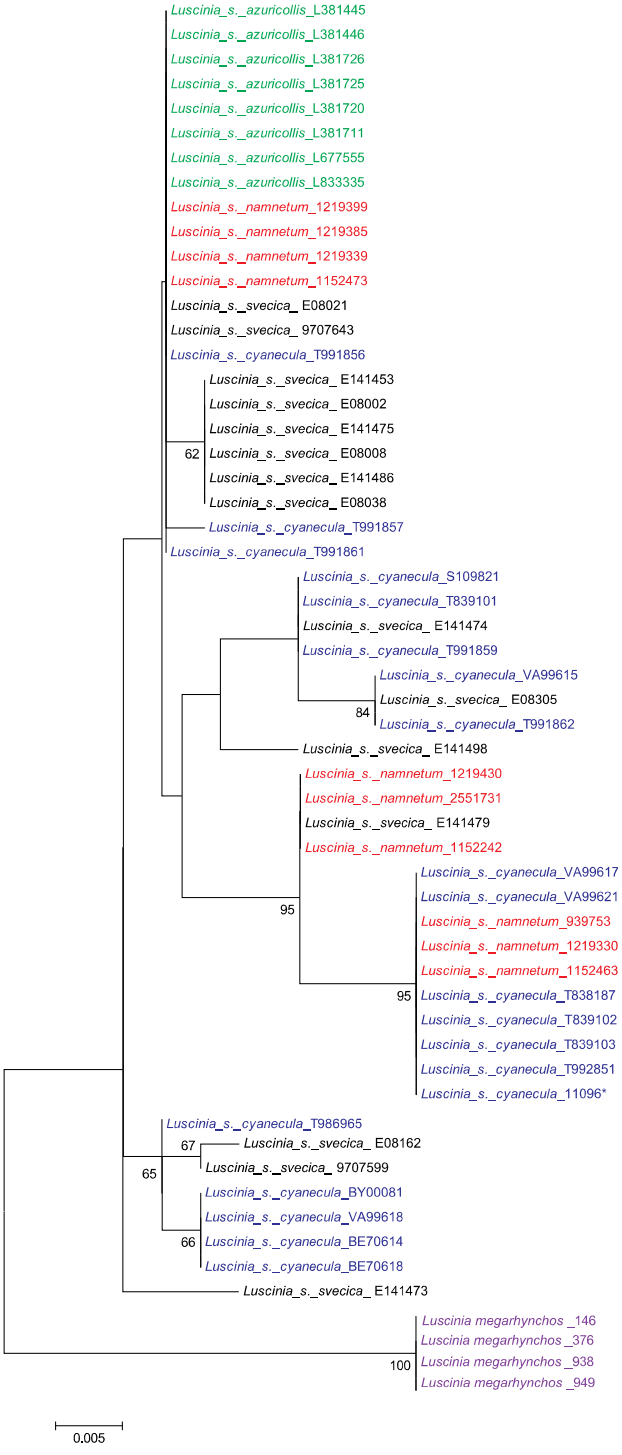
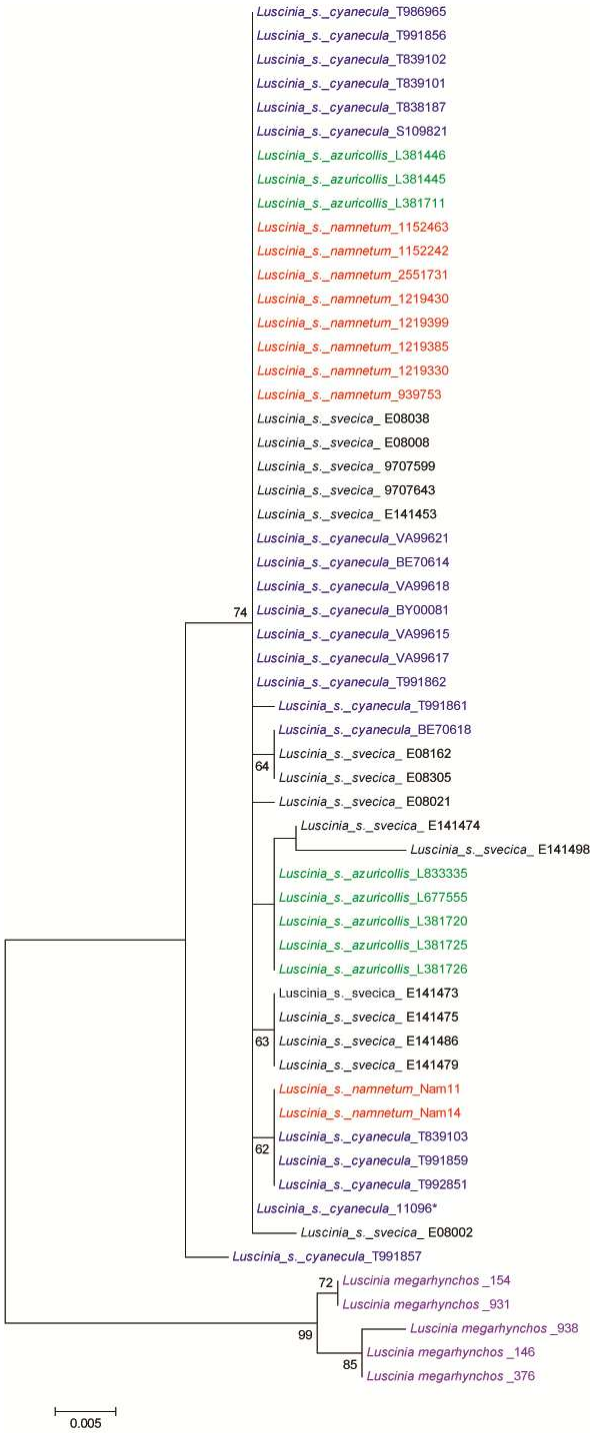

Luscinia_s._azuricollis_L381445
Luscinia_s._azuricollis_L381446
Luscinia_s._azuricollis_L381726
Luscinia_s._azuricollis_L381725
Luscinia_s._azuricollis_L381720
Luscinia_s._azuricollis_L381711
Luscinia_s._azuricollis_L677555
Luscinia_s._azuricollis_L833335
Luscinia_s._namnetum_1219399
Luscinia_s._namnetum_1219385
Luscinia_s._namnetum_1219339
Luscinia_s._namnetum_1152473
Luscinia_s._svecica_ E08021
Luscinia_s._svecica_ 9707643
Luscinia_s._cyanecula_T991856
Luscinia_s._svecica_ E141453
Luscinia_s._svecica_ E08002
Luscinia_s._svecica_ E141475
62
Luscinia_s._svecica_ E08008
Luscinia_s._svecica_ E141486
Luscinia_s._svecica_ E08038
Luscinia_s._cyanecula_T991857
Luscinia_s._cyanecula_T991861
Luscinia_s._cyanecula_S109821
Luscinia_s._cyanecula_T839101
Luscinia_s._svecica_ E141474
Luscinia_s._cyanecula_T991859
Luscinia_s._cyanecula_VA99615
Luscinia_s._svecica_ E08305
84
Luscinia_s._cyanecula_T991862
Luscinia_s._svecica_ E141498
Luscinia_s._namnetum_1219430
Luscinia_s._namnetum_2551731
Luscinia_s._svecica_ E141479
Luscinia_s._namnetum_1152242
Luscinia_s._cyanecula_VA99617
Luscinia_s._cyanecula_VA99621
95
Luscinia_s._namnetum_939753
Luscinia_s._namnetum_1219330
Luscinia_s._namnetum_1152463
95
Luscinia_s._cyanecula_T838187
Luscinia_s._cyanecula_T839102
Luscinia_s._cyanecula_T839103
Luscinia_s._cyanecula_T992851
Luscinia_s._cyanecula_11096*
Luscinia_s._cyanecula_T986965
67
Luscinia_s._svecica_ E08162
Luscinia_s._svecica_ 9707599
65
Luscinia_s._cyanecula_BY00081
Luscinia_s._cyanecula_VA99618
66
Luscinia_s._cyanecula_BE70614
Luscinia_s._cyanecula_BE70618
Luscinia_s._svecica_ E141473
Luscinia megarhynchos _146
Luscinia megarhynchos _376
100
Luscinia megarhynchos _938
Luscinia megarhynchos _949
0.005

Figure S2


Luscinia_s._cyanecula_T986965
Luscinia_s._cyanecula_T991856
Luscinia_s._cyanecula_T839102
Luscinia_s._cyanecula_T839101
Luscinia_s._cyanecula_T838187
Luscinia_s._cyanecula_S109821
Luscinia_s._azuricollis_L381446
Luscinia_s._azuricollis_L381445
Luscinia_s._azuricollis_L381711
Luscinia_s._namnetum_1152463
Luscinia_s._namnetum_1152242
Luscinia_s._namnetum_2551731
Luscinia_s._namnetum_1219430
Luscinia_s._namnetum_1219399
Luscinia_s._namnetum_1219385
Luscinia_s._namnetum_1219330
Luscinia_s._namnetum_939753
Luscinia_s._svecica_ E08038
Luscinia_s._svecica_ E08008
Luscinia_s._svecica_ 9707599
Luscinia_s._svecica_ 9707643
Luscinia_s._svecica_ E141453
Luscinia_s._cyanecula_VA99621
Luscinia_s._cyanecula_BE70614
Luscinia_s._cyanecula_VA99618
74
Luscinia_s._cyanecula_BY00081
Luscinia_s._cyanecula_VA99615
Luscinia_s._cyanecula_VA99617
Luscinia_s._cyanecula_T991862
Luscinia_s._cyanecula_T991861
Luscinia_s._cyanecula_BE70618
64
Luscinia_s._svecica_ E08162
Luscinia_s._svecica_ E08305
Luscinia_s._svecica_ E08021
Luscinia_s._svecica_ E141474
Luscinia_s._svecica_ E141498
Luscinia_s._azuricollis_L833335
Luscinia_s._azuricollis_L677555
Luscinia_s._azuricollis_L381720
Luscinia_s._azuricollis_L381725
Luscinia_s._azuricollis_L381726
Luscinia_s._svecica_ E141473
Luscinia_s._svecica_ E141475
63
Luscinia_s._svecica_ E141486
Luscinia_s._svecica_ E141479
Luscinia_s._namnetum_Nam11
Luscinia_s._namnetum_Nam14
Luscinia_s._cyanecula_T839103
62
Luscinia_s._cyanecula_T991859
Luscinia_s._cyanecula_T992851
Luscinia_s._cyanecula_11096*
Luscinia_s._svecica_ E08002
Luscinia_s._cyanecula_T991857
72
Luscinia megarhynchos _154
Luscinia megarhynchos _931
Luscinia megarhynchos _938
99
Luscinia megarhynchos _146
85
Luscinia megarhynchos _376
0.005

IV

Ecology and Evolution

Open Access

Increased divergence but reduced variation on the Z chromosome relative to autosomes in *Ficedula* flycatchers: differential introgression or the faster-Z effect?

Silje Hogner[1,2], Stein A. Sæther[2], Thomas Borge[2], Torbjørn Bruvik[2], Arild Johnsen[1] & Glenn-Peter Sætre[2]

[1]National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172, Blindern, NO-0318 Oslo, Norway.
[2]Department of Biology, Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, P. O. Box 1066, Blindern, N-0316 Oslo, Norway.

Keywords
Allopatry, congeneric birds, faster-Z hypothesis, nuclear introns, Z chromosome polymorphism.

Correspondence
Silje Hogner, National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172, Blindern, NO-0318 Oslo, Norway. Tel: +47 22 85 18 86; Fax: +47 22 85 18 37;
E-mail: silje.hogner@nhm.uio.no

Funded by The Natural History Museum (to S.H. and A.J.), CEES, University of Oslo (to G.P.S., T.B., and S.A.S.), and the Norwegian Research Council (to G.P.S.).

Received: 28 October 2011; Revised: 09 November 2011; Accepted: 23 November 2011

doi: 10.1002/ece3.92

Abstract

Recent multilocus studies of congeneric birds have shown a pattern of elevated interspecific divergence on the Z chromosome compared to the autosomes. In contrast, intraspecifically, birds exhibit less polymorphism on the Z chromosome relative to the autosomes. We show that the four black-and-white *Ficedula* flycatcher species show greater genetic divergence on the Z chromosome than on the autosomes, and that the ratios of intraspecific polymorphism at Z-linked versus autosomal markers are below the neutral expectation of 75%. In all species pairs, we found more fixed substitutions and fewer shared polymorphisms on the Z chromosome than on the autosomes. Finally, using isolation with migration (IMa) models we estimated gene flow among the four closely related flycatcher species. The results suggest that different pattern of evolution of Z chromosomes and autosomes is best explained by the faster-Z hypothesis, since the estimated long-term gene flow parameters were close to zero in all comparisons.

Introduction

When species diverge from each other, they are expected to gradually lose shared polymorphisms and accumulate fixed substitutions, due to random genetic drift and diversifying selection. Even though levels of polymorphism and divergence are expected to correlate across loci, divergence rates may differ between different parts of the genome. For example, loci under selection and linked sites will show different patterns of variation compared to those evolving neutrally (Nachman 1997; Wang et al. 1997; Fay and Wu 2000; Schlötterer 2003). Also demographic events can cause deviations from patterns expected under neutrality (Kreitman 2000). For example, a reduced population size would lead to a reduction of genetic variation, loss of alleles (especially rare ones), and random changes in allele frequencies (Frankham 1996). One useful approach for separating demographic processes and selection is to analyze patterns of divergence and polymorphism across several unlinked loci. Selection will typically only affect the target loci and closely linked regions whereas demographic processes will affect the whole genome.

Several studies have shown that the macro sex chromosome (X and Z in male and female heterogametic taxa, respectively) play an important role in the evolution of reproductive isolation between closely related species (Presgraves 2008; Qvarnström and Bailey 2009). Hybrids of the heterogametic sex typically suffer greater fitness reduction than homogametic hybrids (Haldane 1922), probably mainly because recessive alleles that are incompatible with heterospecific alleles at other loci get exposed to selection in the heterogametic

sex but stay masked by dominance in the homogametic sex (Turelli and Orr 1995; Orr 1997). Also prezygotic barriers appear often to be controlled by sex-linked genes (reviewed by Qvarnström and Bailey 2009). For instance, studies of the pied flycatcher *Ficedula hypoleuca* and the collared flycatcher *F. albicollis* have shown that Z-linked genes control both male secondary sexual traits and female mate preferences (Sætre et al. 2003; Sæther et al. 2007).

The effective population size of Z-linked loci is ideally 0.75 of that of autosomal loci since females only have one Z-chromosome. Hence, at a balanced sex ratio and equal mutation rates the neutral expectation is that the nucleotide variation at Z-linked genes would be 3:4 of that at autosomal loci (Ellegren 2009). However, several studies of birds have reported much lower ratios than 0.75, suggesting that additional forces are reducing variation on the Z-chromosome relative to autosomes (Berlin and Ellegren 2004; Borge et al. 2005b; Storchova et al. 2010; Backström and Väli 2011; Elgvin et al. 2011). At the same time, a pattern of elevated interspecific divergence on the Z chromosome compared to the autosomes has also been found in birds (e.g., Borge et al., 2005b; Storchova et al. 2010; Elgvin et al. 2011). One explanation for this apparent nonneutral pattern is the faster-Z hypothesis (Charlesworth et al. 1987). Faster adaptive evolution on the Z is expected because (partially) recessive beneficial mutations are not masked by dominance in the heterogametic sex. Likewise, (partially) recessive deleterious mutations would be more effectively purged on the Z compared to autosomes due to hemizygous exposure. Associated selective sweeps on the Z chromosome would contribute to further reduce intraspecific polymorphism (Charlesworth et al. 1987; Borge et al. 2005b). Genetic drift can also contribute to a faster-Z effect because the lower effective population size of the Z chromosome would be associated with increased rates of genetic drift and thus an increased fixation rate of mildly deleterious mutations (Charlesworth et al. 1987; Mank et al. 2010).

A second hypothesis, here termed the differential introgression hypothesis, is that the accumulation of incompatibilities on the Z-chromosome may reduce the rate of introgression of Z-linked compared to autosomal genes and essentially produce the same pattern as predicted by the faster-Z hypothesis (Carling et al. 2010; Storchova et al. 2010; Backström and Väli 2011). The two hypotheses are not mutually exclusive, however. For instance, a faster-Z effect may speed up divergence and hence contribute to the accumulation of sex-linked incompatibilities that would reduce Z-linked introgression (e.g., Elgvin et al. 2011).

In this study, we test between the two hypotheses by analyzing pattern of polymorphism and divergence on Z-linked and autosomal loci in all the four species of the old world black-and-white flycatcher complex: the pied (*F. hypoleuca*), collared (*F. albicollis*), Atlas (*F. speculigera*), and semicollared flycatcher (*F. semitorquata*) (Fig. 1). The pied and collared flycatchers have earlier been investigated in some detail with respect to the role of selection in speciation (see Qvarnström et al. 2010; Sætre and Sæther 2010 for recent reviews). A moderate level of gene flow has been observed at autosomal loci between the collared and the pied flycatchers living in sympatry, while introgression on the Z chromosome is apparently absent (Sætre et al. 2003; Borge et al. 2005a; Backström et al. 2010). In a study of allopatric pied and collared flycatchers (from Spain and Italy, respectively), Borge et al. (2005b) reported significantly reduced levels of genetic variation but elevated rate of divergence at Z-linked loci in both species. However, Borge et al. (2005b) were unable to discriminate between the faster-Z effect and historical autosomal introgression as explanations for the pattern. At present, only the pied and the collared flycatcher have overlapping breeding ranges and hybridize. Even if some of the other species may also have hybridized in the past, we consider it unlikely that all the four flycatcher species have exchanged genes to the same extent. Hence, including more species is likely to help disentangling the relative importance of differential introgression and faster-Z in shaping the genomes of these birds.

To critically test between the faster-Z and the differential introgression hypotheses, we use isolation with migration model analysis (IMa) to estimate key demographic parameters, including effective population sizes, divergence times, and levels of gene flow between the species pairs at both the Z-linked and autosomal datasets. From the differential introgression hypothesis, we predict higher estimates of gene flow between the species at the autosomal compared to the Z-linked dataset, and that the amount of estimated gene flow on Z should be negatively associated with the degree of elevated divergence and reduced polymorphism on the Z relative to the autosomes among the different species pairs. From the faster-Z hypothesis, we predict no difference in amount of gene flow between autosomal and Z-linked loci.

Materials and Methods

Samples

Adult breeding birds of the pied, collared, semicollared, and Atlas flycatchers were caught at their respective breeding grounds (Fig. 2). The Atlas flycatchers are from near Azrou, Morocco ($N = 15$ males, 33°26′N, 5°13′W), the collared flycatchers are from near Pescasseroli, Italy ($N = 16$ males, 41°48′N, 13°47′E) and from Pilis Mts, Hungary ($N = 16$ males, 47°43′N, 19°01′E), the pied flycatchers are from near Oslo, Norway ($N = 16$ males, 59°59′N, 10°46′E) and from La Granja, Spain ($N = 14$ males and two females, 40°14′N, 55°93′W), and the semicollared flycatchers are from Kamchia, Bulgaria ($N = 12$ males and three females, 42°53′N, 26°58′E). All birds were caught using mist nets and playback. Additionally, one male red-breasted flycatcher caught

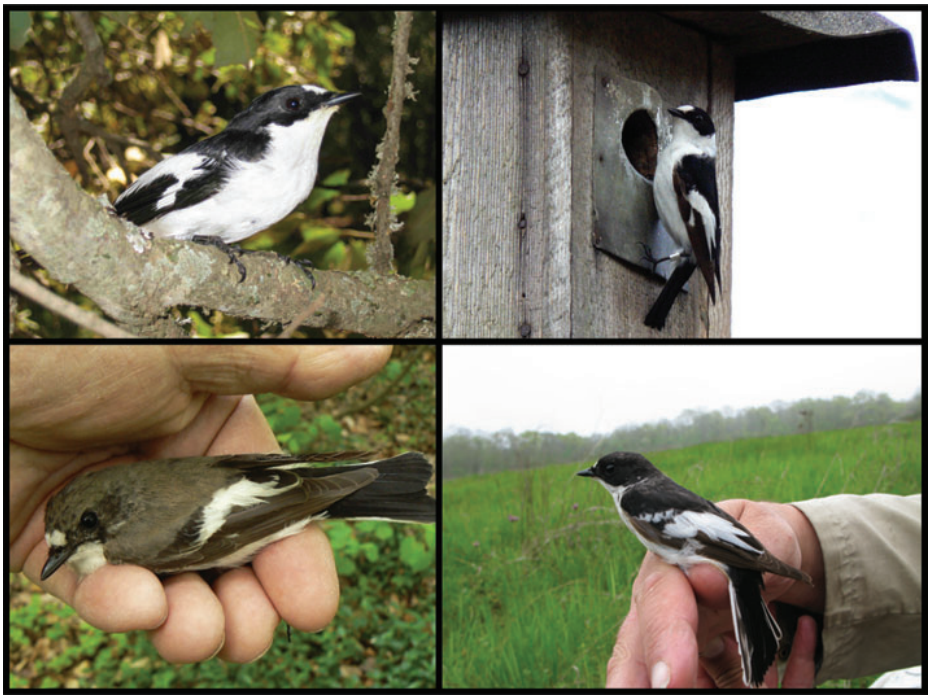


Figure 1. Males of the four study species. Top left: Atlas flycatcher *Ficedula speculigera* (photo: Gunilla Andersson), top right: collared flycatcher *F. albicollis* (photo: Miroslav Král), bottom left: pied flycatcher *F. hypoleuca* (photo: Miroslav Král), bottom right: semicollared flycatcher *F. semitorquata* (photo: Silje Hogner).

in Northern Moravia, Czech Republic, is included in the analysis to serve as an outgroup.

DNA preparation

Approximately 25 μl of blood were collected from each individual by brachial venipuncture, and suspended in 1 mL Queen's lysis buffer (Seutin et al. 1991). DNA extractions were done using two different kits, QIAamp DNA Blood Mini Kit (Qiagen AB, Sollentuna, Sweden) and eZNA Blood DNA Kit (Omega Bio-Tek, Norcross, Georgia, USA), following the manufacturers protocols for each of the two kits.

Several primer pairs that have been tested on *Ficedula* flycatchers were available from previous studies (Primmer et al. 2002; Borge et al. 2005b). These primers were designed as described by Primmer et al. (2002), using chicken (*Gallus gallus*) sequences, available from GenBank, as templates. The primers were designed in exon sequences flanking introns of sizes appropriate for direct sequencing of both autosomal and Z-linked genes. New flycatcher-specific primers were designed for long introns when the amplification successes for these were variable (Borge et al. 2005b). In this study, we used primers that had high amplification success in earlier studies on the pied and collared flycatcher (Primmer et al. 2002; Sætre et al. 2003; Borge et al. 2005b) and that yielded high-quality sequences also in the other two species (Table 1).

The introns were amplified in PCR reaction volumes of 10 μl, containing dH_2O, 1× PCR buffer II (Applied Biosystems), 1.5 mM magnesium, 0.2 mM dNTP (ABgene, Epsom, UK), 0.5 mM forward and reverse primer, 3% Dimethyl sulfoxide (DMSO), 0.25 U AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California, USA), and approximately 50 ng DNA template. The amplifications were run on a DNA Engine Tetrad 2 (MJ Research, Waterton, MA). The following profile was used: 95°C for 1 min, 94°C for 30 sec, primer-specific annealing temperature (see Table 1) (55–60°C) for 30 sec, 72°C for 1 min, then the second and third step another 34–39 cycles before the last step, 72°C for 10 min. Three milliliters of PCR product was electrophoresed in 1% Tris/Borate/EDTA (TBE) agarose to confirm amplification success and to exclude any contamination.

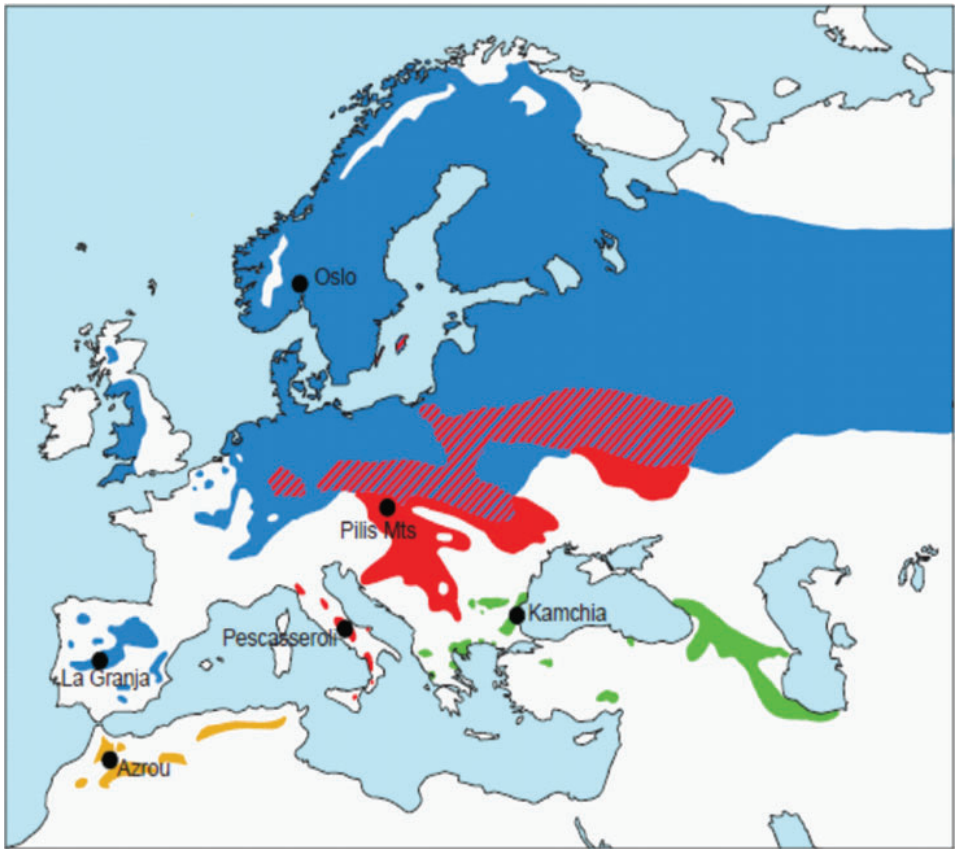


Figure 2. Breeding distribution of the Atlas flycatcher *Ficedula speculigera* (yellow), collared flycatcher *F. albicollis* (red), pied flycatcher *F. hypoleuca* (blue), and semicollared flycatcher *F semitorquata* (green). The hatched area indicates areas of distributional overlap between the pied and collared flycatcher. Sampling localities are indicated with black dots.

The remaining PCR product was purified by digesting unincorporated nucleotides and primers using diluted (1:9) ExoSap-It (United States Biochemical Cleveland) run at 37°C for 45 min followed by 80°C for 15 min to inactivate the enzyme. The PCR products were then sequenced using BigDye Terminator sequencing buffer and v 3.1 Cycle Sequencing kit (Applied Biosystems). The sequences were aligned and edited using ClustalW in the program Mega 4.0.2 (Tamura et al. 2007) or in Sequencher 4.1 (Gene Codes, Ann Arbor, MI) and modified manually. Each base was called, using at least single-fold coverage sequencing reads for each strand. All sequences for each locus were adjusted to the same length as the shortest sequence of that locus for comparison. We analyzed intronic sequences from five Z-linked and six autosomal loci (accession numbers: JN995666-JN996468).

Previous reviewers have suggested that the sample sizes in this study are not sufficient to distinguish between the faster-Z and the differential introgression hypotheses. We disagree for three main reasons. First, previous studies utilizing a similar number of loci and individuals have found significant differences in level of polymorphisms and divergence between Z-linked and autosomal genes (e.g., Storchova et al. 2010; Elgvin et al. 2011). Second, our study is unique in the sense that we compare several geographically dispersed populations/species that differ in the likelihood of having experienced episodes of introgressive hybridization. Although both hypotheses predict reduced polymorphism and increased divergence of Z-linked compared to autosomal loci, they differ with respect to the heterogeneity of this signal. The faster-Z hypothesis predicts that all pairs of populations will be

Table 1. Primer information and amplification conditions.

Locus	C[1]	Primer sequence (5′-3′)[2]	Annealing temperature
Acly-16	A	F: ACCATGAATTATCCCCAGGTGAG R: CAAAACCATTGGTACCCCACAG	55
Alas1-8	A	F: CCGAGTCACATCATTCCCGT R: AGCAGCATCTGCAACCTGAA	55
Fas-y	A	F: TGAAGAAGGTCTGGGTGGAGA R: CTCCAATAAGGTGCGGTGA	50
Rho-1	A	F: CATCGAGGGCTTCTTTGCC R: TTTAGACACACAATTTCTATTTAACACCTGT	55
Rpl30-3	A	F: CCAAGTTGGTCATCCTAGCCA R: GCCACTATAATGATGGACACCAGTC	60
Tgfb2-5	A	F: TGCCTGCCATACATCCAGTG R: TGCTTGCTTCCTGAATGATCCT	55
Aldob-6	Z	F: AGACCATGATCTCCAGCGCT R: CCTTCCAGGTAGACATGATG	55
Brm-12	Z	F: CCCTATCTCATCATTGTTCC R: CACAGAAGGAGCCCATTTGT	50
Chdz_15	Z	F: TAGAGAGATTGAGAACTACAGT R: GACATCCTGGCAGAGTATCT	52
Chdz_18	Z	F: TACATACAGGCTCTACTCCT R: CCCCTTCAGGTTCTTTAAAA	58
Ghr5_1	Z	F: GCTTCCATTATGTATCTTACC R: TTTGGCTTCTAGAGTTTTGCA	55
GHR5_2	Z	F: ATGTTATTGCTTGTTCAGAGTG R: GAGTATTTGGAATAAAACAGCC	58
Vldlr-8	Z	F: GTTATTGGCTATGAATGTGA R: GTTGATACAGATTTGGCTAC	54
Vldlr-9	Z	F: AAGTGTGAATGTAGCCGTGG R: TCGGTTGGTGAAAATCAGAC	54
Vldlr-12	Z	F: GTTCCTTCCTCATCCTCTTG R: ATAGACTGCCTCGTTCTCTC	55

[1]Chromosome class: A = autosomal; Z = Z-linked.
[2]Borge et al. (2005).

similarly affected, whereas the differential introgression hypothesis would only affect population pairs that have hybridized extensively after partial reproductive isolation has developed. Finally, our sample sizes allow good estimates of historic gene flow between several population pairs using a coalescent simulation framework (see below).

Cloning

Sequences with difficult gaps were cloned in order to get the respective haplotypes. All cloning was performed using the Zero Blunt TOPO PCR cloning kit (Invitrogen) and transformed into *Escherichia coli* DH5α chemically competent cells as recommended by Invitrogen (Invitrogen, Carlsbad, California, USA). The transformed cells were then grown on LB agar containing Kanamycin (100 μg/mL) as a selection marker. Eight positive colonies from each cloned individual were picked with sterile toothpicks and diluted with 6 μl dH_2O and used directly as PCR template. Amplification and purification was performed as described for genomic DNA and standard M13 primers were used.

Polymorphism and divergence

The program DnaSP 4.0 (Librado and Rozas 2009) was used to analyze polymorphism. Two common measures of nucleotide polymorphism were calculated: π, the average number of nucleotide differences per site between two sequences (Nei 1987) and θ_W, the population mutation parameter estimated from the number of segregating sites in the aligned sample of sequences. $\theta = 4N_e\mu$ for autosomal loci and $3N_e\mu$ for Z-linked loci, in which N_e is the effective population size and μ is the neutral mutation rate (Nei 1987). The π and θ_W parameters were estimated with standard deviations in DnaSP.

We calculated the Z:autosomal ratio (Z_θ:A_θ) of average pairwise sequence differences. This was done by dividing the average pairwise sequence difference per nucleotide for each Z-linked locus by the average pairwise sequence difference per nucleotide for all autosomal loci.

Table 2. Polymorphism summaries of Z-linked genes in the four black-and-white *Ficedula* flycatcher species.

Gene	Species[1]	N[2]	L[3]	K[4]	S[5]	s[6]	Π[7]	θW[8]	Tajima's D[9]
Aldob-6	A-Mar	30	437	10	0	0	0	0	–
	C-It	30	437	9	1	0	0.0004	0.0006	−0.409
	P-Spa	31	437	9	3	1	0.0006	0.0017	0.179
	C-Hun	32	437	9	1	0	0.0004	0.0006	−0.448
	P-Nor	32	437	9	0	0	0	0	–
	S-Bul	27	437	10	1	1	0.0002	0.0006	−1.154
Brm-12	A-Mar	28	1439	30	20	11	0.0017	0.0036	−1.897
	C-It	28	1438	33	11	4	0.0014	0.0002	−0.910
	P-Spa	24	1439	34	12	1	0.0030	0.0022	1.118
	C-Hun	30	1439	30	26	11	0.0037	0.0046	−0.704
	P-Nor	26	1438	32	23	8	0.0033	0.0042	−0.773
	S-Bul	25	1434	23	15	8	0.0030	0.0043	−1.077
Chdz	A-Mar	30	638	14	1	0	0.0005	0.0004	0.216
	C-It	29	638	13	1	1	0.0001	0.0004	−1.149
	P-Spa	31	638	14	3	1	0.0006	0.0012	−1.183
	C-Hun	30	638	13	4	2	0.0006	0.0016	−1.574
	P-Nor	32	638	14	1	0	0.0004	0.0004	0.147
	S-Bul	27	638	13	1	0	0.0004	0.0004	0.017
Ghr	A-Mar	30	548	9	3	0	0.0021	0.0014	1.147
	C-It	30	555	9	4	3	0.0010	0.0018	−0.796
	P-Spa	31	555	10	4	1	0.0008	0.0018	−1.430
	C-Hun	32	555	9	4	0	0.0014	0.0018	−0.541
	P-Nor	32	555	10	6	2	0.0013	0.0027	−1.456
	S-Bul	27	555	10	5	2	0.0020	0.0023	−0.452
Vldlr	A-Mar	30	562	5	9	2	0.0030	0.0040	−0.771
	C-It	30	562	4	12	3	0.0062	0.0054	0.464
	P-Spa	28	562	4	3	0	0.0007	0.0014	−1.115
	C-Hun	32	558	4	13	6	0.0049	0.0058	−0.481
	P-Nor	30	562	4	4	2	0.0011	0.0018	−0.960
	S-Bul	27	562	4	11	8	0.0026	0.0051	−1.598
Total Z	**A-Mar**	**30**	**3624**	**68**	**33**	**13**	**0.0014**	**0.0015**	**−0.161**
	C-It	**30**	**3631**	**68**	**29**	**11**	**0.0020**	**0.0021**	**−0.102**
	P-Spa	**31**	**3631**	**71**	**25**	**4**	**0.0010**	**0.0015**	**−1.113**
	C-Hun	**32**	**3641**	**65**	**48**	**19**	**0.0024**	**0.0029**	**−0.580**
	P-Nor	**32**	**3641**	**69**	**34**	**12**	**0.0009**	**0.0012**	**−0.736**
	S-Bul	**27**	**3626**	**60**	**33**	**19**	**0.0013**	**0.0021**	**−1.338**
Acly-16	A-Mar	30	358	5	0	0	0	0	–
	C-It	30	358	5	1	1	0.0002	0.0007	−1.147
	P-Spa	30	358	5	2	1	0.0006	0.0014	−1.256
	C-Hun	32	358	5	2	2	0.0004	0.0014	−1.504
	P-Nor	32	358	5	0	0	0	0	–
	S-Bul	30	358	5	1	1	0.0002	0.0007	−1.147
Alas1-8	A-Mar	30	290	12	7	2	0.0051	0.0061	−0.479
	C-It	30	290	14	9	4	0.0054	0.0078	−0.957
	P-Spa	34	290	15	6	2	0.0027	0.0051	−1.302
	C-Hun	32	290	13	9	3	0.0062	0.0077	−0.620
	P-Nor	32	290	15	7	3	0.0023	0.0060	−1.786
	S-Bul	30	288	13	9	2	0.0083	0.0079	0.144
Fas-y	A-Mar	26	551	10	0	0	0	0	–
	C-It	26	552	11	7	3	0.0018	0.0033	−1.377
	P-Spa	34	551	10	1	1	0.0001	0.0004	−1.138
	C-Hun	32	550	11	1	1	0.0001	0.0005	−1.142
	P-Nor	30	551	10	2	1	0.0009	0.0009	−0.136
	S-Bul	30	551	9	6	1	0.0013	0.0028	−1.515

Table 2. Continued.

Gene	Species[1]	N[2]	L[3]	K[4]	S[5]	s[6]	Π[7]	θw[8]	Tajima's D[9]
Rho-1	A-Mar	30	372	8	8	2	0.0049	0.0054	−0.306
	C-It	30	372	7	6	1	0.0049	0.0041	0.577
	P-Spa	30	372	7	8	1	0.0072	0.0054	0.970
	C-Hun	26	371	7	11	6	0.0060	0.0078	−1.002
	P-Nor	32	372	7	6	1	0.0057	0.0047	0.644
	S-Bul	30	372	6	12	4	0.0070	0.0081	−0.450
Rpl30-3	A-Mar	30	983	17	22	7	0.0058	0.0057	0.107
	C-It	30	983	16	21	2	0.0053	0.0054	−0.051
	P-Spa	30	983	17	14	7	0.0021	0.0036	−1.392
	C-Hun	32	982	20	23	13	0.0043	0.0058	−0.893
	P-Nor	28	983	17	21	2	0.0083	0.0055	1.806
	S-Bul	26	983	21	15	3	0.0055	0.0040	1.295
Tgfb2-5	A-Mar	30	402	5	6	2	0.0042	0.0038	0.308
	C-It	30	402	5	6	4	0.0021	0.0038	−1.277
	P-Spa	34	402	5	3	0	0.0021	0.0018	0.289
	C-Hun	32	402	5	11	4	0.0047	0.0068	−1.354
	P-Nor	32	402	5	7	4	0.0030	0.0043	−0.875
	S-Bul	30	401	5	6	3	0.0043	0.0038	−0.075
Total A	**A-Mar**	**30**	**2956**	**57**	**43**	**13**	**0.0045**	**0.0045**	**−0.054**
	C-It	**30**	**2957**	**58**	**50**	**15**	**0.0040**	**0.0045**	**−0.455**
	P-Spa	**34**	**2956**	**59**	**34**	**12**	**0.0013**	**0.0020**	**−0.986**
	C-Hun	**32**	**2958**	**61**	**57**	**29**	**0.0031**	**0.0044**	**−1.177**
	P-Nor	**32**	**2975**	**59**	**43**	**11**	**0.0028**	**0.0037**	**−0.799**
	S-Bul	**30**	**2956**	**59**	**49**	**14**	**0.0038**	**0.0044**	**−0.557**

[1]A-Mar = Atlas flycatcher from Morocco; C-It = collared flycatcher from Italy; P-Spa = pied flycatcher from Spain; C-Hun = collared flycatcher from Hungary; P-Nor = pied flycatcher from Norway; S-Bul = semicollared flycatcher from Bulgaria.
[2]Number of sites surveyed.
[3]Sequence length.
[4]Number of divergent sites with outgroup.
[5]Number of segregating sites.
[6]Number of singleton sites.
[7]Average pairwise sequence difference per nucleotide (Nei 1987).
[8]Expected heterozygosity per nucleotide (Watterson 1975).
[9]None of the D-values significant after correcting for multiple tests.

Polymorphisms were divided into four categories for each of the species pairs: variable sites exclusive to one of the species, shared polymorphisms, fixed differences, and the average number of pairwise differences.

To test if the levels of polymorphism and divergence were correlated between loci and species as predicted under neutrality, multilocus Hudson–Kreitmann–Aguade (HKA) tests (Hudson et al. 1987) were run for all species pairs for both Z-linked and autosomal loci using the online software of Jody Hey's lab (http://lifesci.rutgers.edu/~heylab). Deviations from the expected relationship may indicate that selection has affected genetic variation. The polymorphism and divergence data from Table 2 and Table S1 were used for these tests. For all species pairs, 10,000 coalescent simulations were run to assess significance.

DnaSP was used to compute Tajima's *D* (Tajima 1989). This neutrality test is based on the allele frequency spectrum. It can be used to infer previous evolutionary and demographic events that the populations have experienced. Negative values of Tajima's *D* reflect an excess of rare polymorphisms in the populations, while positive values indicates an excess of intermediate-frequency alleles. An excess of rare alleles might result from positive selection or an increase in population size, whereas an excess of intermediate-frequency alleles might result from balancing selection or a population bottleneck (Akey et al. 2004).

Population divergence

For sequences that contained more than one heterozygous site, and in which cloning was not performed, haplotypes were inferred using the programme Phase version 2.1.1 (Stephens and Donnelly 2003). Each run had the following values set, iterations (10,000), thinning interval (1), and burn-in (1000). Harrigan et al. (2008) showed that haplo-

types with a PHASE probability greater than 0.5 are reliable. Consequently, we used 0.5 as a lower limit.

Analysis of molecular variance (AMOVA) was run using Arlequin v 3.5 (Excoffier et al. 2005) to examine the genetic structure of the populations. The variance was partitioned into four: between the species, between individuals within the species, among individuals within populations, and within the different individuals. Pairwise species differentiation was estimated using F_{ST} (Weir and Cockerham 1984), with default settings in the population comparisons. These F_{ST} values can be used as short-term genetic distances between populations. The null distribution of pairwise F_{ST} values is obtained by permutating haplotypes between populations under the hypothesis of no differences between populations. The *P*-value of the test is given as the proportions of simulations giving an F_{ST} value larger or equal to the observed one.

IMa model

Recombination was calculated as the minimum number of recombination events (RM) using the four-gamete test in DnaSP (Hudson and Kaplan 1985). The sites with recombination were excluded from the coalescent analysis.

In order to estimate the timing, magnitude of divergence and gene flow between the four species, a coalescent framework, the IMa (Nielsen and Wakeley 2001; Hey and Nielsen 2007) model was used. We included one population from each of the four species. We chose the Norwegian population of the pied flycatcher because previous analyses had indicated inbreeding effects (violations of the neutral model) in the alternative Spanish population (Haavie et al. 2000). Further, we included the Italian population of the collared flycatcher because the alternative Hungarian population is located close to the pied × collared flycatcher hybrid zone, and hence, could have been affected by recent introgression (Borge et al. 2005a).

Initial IMa runs were conducted using wide priors set as recommended in the IMa manual. Two final sets of runs were conducted, one with only one chain (run multiple times with different random seed numbers) and one with 10 chains and 10 swapping events. First, the multiple final runs were conducted with a length between 8.0×10^7 and 9.0×10^8 generations, where the first 30% were discarded as burn-in. Second, a run contained 10 chains and 10 chain swap attempts per step, with a burn-in period of 10–15% and a run length between 2.0×10^7 and 3.0×10^7 generations. No runs were stopped before the effective sample size values had exceeded at least 200. Two separate IMa analyses were run for each species pairs: (1) the autosomal dataset containing six autosomal loci and (2) the Z-linked dataset containing five Z-linked loci. Two independent runs, with different random seed numbers, were conducted per comparison.

In order to convert the parameter estimates into demographic quantities, we used a neutral mutation rate of 1.35×10^{-9} substitutions per site per year for autosomal genes (Ellegren 2007) and 1.45×10^{-9} substitutions per site per year for Z-linked genes (Axelsson et al. 2004; Ellegren 2007). Using these mutation rates and setting the generation time to one year, we calculated the geometric mean of mutation rates per locus for the different datasets. There was good agreement between the independent runs and therefore we only report the results from the longest independent run.

Results

Intraspecific polymorphism

We found that the frequency of polymorphic sites was quite heterogeneous among loci and somewhat higher overall for the autosomal loci than for the Z-linked loci (Table 2; Fig. 3). The total frequency of polymorphic sites was significantly higher at the autosomal loci than the Z-linked ones in the Atlas flycatcher ($P = 0.048$), the Italian collared flycatcher ($P = 0.0013$), The Spanish pied flycatcher ($P = 0.050$), and the semicollared flycatcher ($P = 0.0072$), and nearly significantly higher also in the Hungarian collared flycatcher ($P = 0.060$) and the Norwegian pied flycatcher ($P = 0.064$) (Fisher's exact tests). Combining the *P*-values above in a Fisher's combined probability test yields a highly significant test statistic ($\chi^2 = 46.35$, df = 12, $P < 0.0001$). Hence, there is an overall reduction in the frequency of polymorphic sites at the Z-linked loci in the dataset. For all species except the Spanish pied flycatchers, the Z_θ:A_θ ratios were below the expected value of 0.75 (Fig. 4). The ratio was significantly lower than 0.75 in the Atlas flycatcher ($t = 2.72$, df = 10, $P = 0.022$, one-sample *t*-test) and the Italian collared flycatcher ($t = 2.23$, df = 10, $P = 0.050$) and nearly so also in the Norwegian pied flycatcher ($t = 1.82$, df = 10, $P = 0.098$). The other ratios did not differ significantly from the expectation of 0.75 ($P > 0.1$, one-sample *t*-tests). However, when combining all Z_θ:A_θ ratios, Fisher's combined *P*-values were significant ($\chi^2 = 25.80$, df = 12, $P = 0.012$).

Fixed and shared polymorphisms between species

For all the species combinations, the average number of shared polymorphisms was higher at the autosomal than the Z-linked loci (Fig. 5). In contrast, the level of fixed differences was in most cases higher at the Z-linked than at the autosomal loci (Fig. 6). The ratio of fixed differences to shared polymorphisms was significantly higher at the Z-linked loci compared to autosomal ones in five of 13 species comparisons (Atlas flycatcher vs. Italian collared flycatcher, Atlas flycatcher vs. Norwegian pied flycatcher, Italian collared flycatcher vs. Norwegian pied flycatcher, Italian collared flycatcher vs. Spanish pied flycatcher, and semicollared fly-

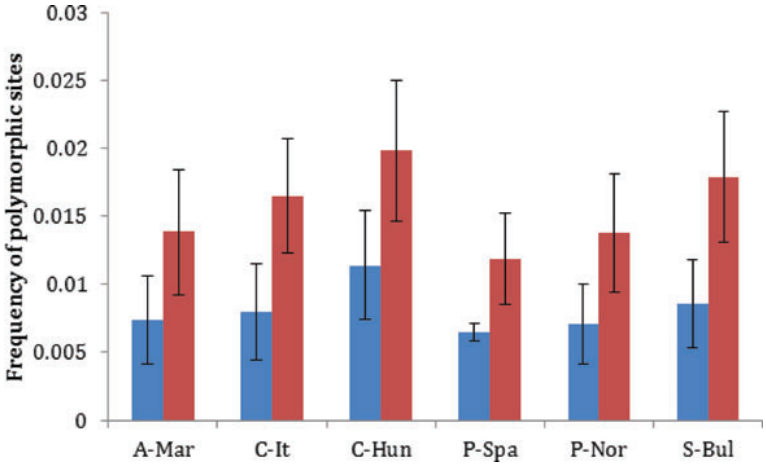


Figure 3. Mean frequency (± SE) of polymorphic sites for Z-linked (blue) and autosomal (red) loci in the six populations. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; C-Hun, collared flycatcher from Hungary; P-Spa, pied flycatcher from Spain; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria.

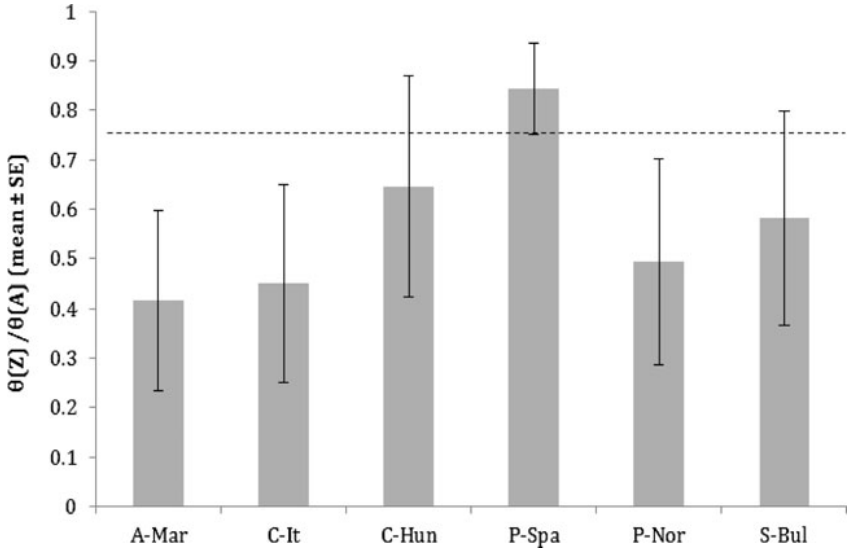


Figure 4. The ratio of Z to autosomal variation measured as the expected heterozygosity for each Z-linked locus (θ(z)) divided by the expected heterozygosity for all autosomal loci (θ(A)). The dashed line indicates the expected value of 0.75. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; C-Hun, collared flycatcher from Hungary; P-Spa, pied flycatcher from Spain; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria.

catcher vs. Norwegian pied flycatcher) (Fisher's exact tests: $P < 0.05$); and nearly significantly higher also between Atlas flycatcher versus Spanish pied flycatcher, Hungarian collared flycatcher versus Norwegian pied flycatcher, and Atlas flycatcher versus semicollared flycatcher (Fisher's exact tests: $0.1 > P > 0.05$). However, when using Bonferroni correction for multiple comparisons and setting alpha to 0.05, only three comparisons remain statistically significant (Atlas flycatcher

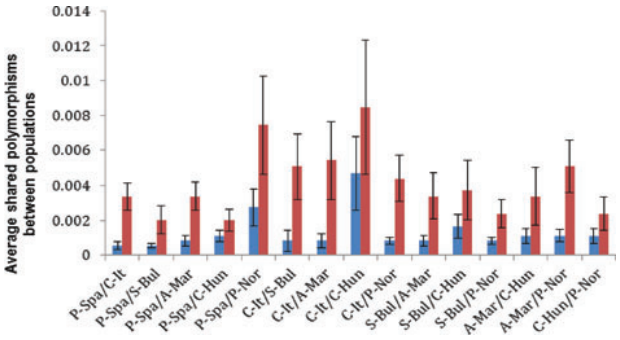


Figure 5. Average number of shared polymorphisms ± SE at Z-linked (blue) and autosomal (red) loci between the different flycatcher populations. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; C-Hun, collared flycatcher from Hungary; P-Spa, pied flycatcher from Spain; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria.

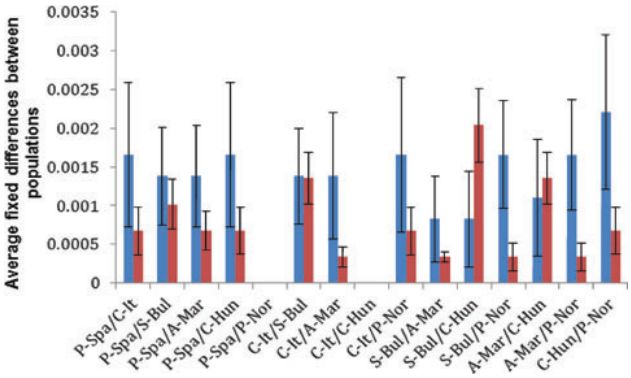


Figure 6. Average number of fixed differences ± SE at Z-linked (blue) and autosomal (red) loci between the different flycatcher populations. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; C-Hun, collared flycatcher from Hungary; P-Spa, pied flycatcher from Spain; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria.

vs. Italian collared flycatcher, Italian collared flycatcher vs. Norwegian pied flycatcher, and Atlas flycatcher vs. Norwegian pied flycatcher).

Tests of neutrality

HKA tests for all population comparisons revealed no significant deviations from expected values at neither the autosomal nor the Z-linked loci (all P-values > 0.17; Table S2). We also performed neutrality test based on the allele frequency distribution, namely Tajima's D for each locus. Again, none of the tests showed any significant deviations from neutrality (Table 2). For all species except the Atlas flycatcher and the Norwegian pied flycatchers, the Tajima's D values were slightly negative at both the Z-linked and the autosomal loci. Hence, the allele frequency spectra closely match the neutral expectations, with a small skew toward rare alleles (and a small skew toward alleles of intermediate frequencies in the Atlas flycatcher and the Norwegian pied flycatchers).

Population divergence

AMOVA for all Z-linked loci combined showed that about 76% of the variation could be explained by differences between species, 7% by variation among individuals within populations, 16% by variation within individuals, and less than 1% from variation among populations within groups. For all autosomal loci combined, 47% of the variation could be explained by differences between species, 5% by variation between individuals within populations, 47% by variation within individuals, and around 1% by variation among populations within groups (Table 3). We also computed F_{ST} values for each of the species pairs (Table 4). All the F_{ST} values were high, quite similar between the different species pairs and significantly larger than zero, all $P < 0.01$. Hence, all the four species are genetically strongly differentiated from each other. Using pairwise comparisons, F_{ST} was significantly higher at the Z-linked dataset compared to the autosomal dataset (Table 2; $t = 7.06$, df $= 14$, $P < 0.0001$, paired t-test). Restricting the test to between-species comparisons further increased the significance ($t = 10.1$, df $= 12$, $P < 0.000001$, paired t-test).

IMa analysis

We found very low estimates of levels of gene flow at both the Z-linked and autosomal dataset in all species pairs (Fig. 7;

Table 3. Analysis of molecular variation (AMOVA) between six populations (four species) of *Ficedula* flycatchers.

Source of variation	df	Sum of squares	Variance components	Percentage of variation	*P* -value
Z-linked loci					
Between species	3	1380.82	5.38	76.41	<0.05
Among populations within species	2	8.77	0.04	0.53	<0.05
Among individuals within populations	170	361.68	0.50	7.16	<0.001
Within individuals	176	197.00	1.12	15.90	<0.001
Total	351	1928.26	7.04		
Autosomal loci					
Between species	3	643.12	2.28	47.09	<0.05
Among populations within species	2	12.45	0.05	1.11	<0.01
Among individuals within populations	182	503.96	0.26	5.30	<0.001
Within individuals	188	424.00	2.26	46.50	<0.001
Total	375	1574.53	4.85		

Table 4. Pairwise F_{ST} values between six populations (four species) of *Ficedula* flycatchers; Atlas flycatcher from Morocco (A-Mar), collared flycatcher from Italy (C-It) and Hungary (C-Hun), pied flycatcher from Spain (P-Spa) and Norway (P-Nor), and semicollared flycatcher from Bulgaria (S-Bul). All values are significant except the one marked with an asterisk.

	P-Spa	P-Nor	C-Hun	C-It	S-Bul	A-Mar
Z-linked loci						
P-Spa	0.000					
P-Nor	0.049	0.000				
C-Hun	0.692	0.699	0.000			
C-It	0.703	0.710	0.016	0.000		
S-Bul	0.760	0.763	0.557	0.578	0.000	
A-Mar	0.851	0.853	0.819	0.827	0.853	0.000
Autosomal loci						
P-Spa	0.000					
P-Nor	0.034	0.000				
C-Hun	0.420	0.438	0.000			
C-It	0.484	0.504	0.020*	0.000		
S-Bul	0.466	0.498	0.412	0.441	0.000	
A-Mar	0.605	0.629	0.374	0.382	0.500	0.000

Table 5). In fact, the posterior probability densities approached zero migration in all runs (Fig. 7). The divergence time estimates suggested more recent splits between the species pairs at the autosomal dataset (range 298,086–756,159 years ago) compared to the Z-linked dataset (range 385,550–1, 245,572 years ago) (Fig. 8; Table 5). The effective population size estimates were high in all four species and quite similar at the autosomal and Z-linked datasets (Fig. 9; Table 5). For most parameters in most runs, we got good convergence and the posterior probability tails were zero or close to zero within the parameter range. For the ones that did not reach zero, the highest and most probable peak was within the prior parameter range.

Discussion

We analyzed polymorphism and divergence at five Z-linked and six autosomal loci in and between the four black-and-white *Ficedula* flycatcher species. The Z-linked markers exhibited reduced levels of polymorphism, yet elevated levels of differentiation between the species pairs compared to the autosomal ones. Such a pattern of elevated divergence and reduced polymorphism has previously been reported also among other closely related bird species (Berlin and Ellegren 2004; Borge et al. 2005b; Storchova et al. 2010; Backström and Väli 2011; Elgvin et al. 2011), and has alternately been attributed to faster adaptive divergence on the Z chromosome (the faster-Z hypothesis) and/or reduced introgression on the Z chromosome due to accumulation of sex-linked incompatibilities (the differential introgression hypothesis). Below, we discuss our results in more detail to try to disentangle which of the hypotheses best explains the observed patterns.

Comparison of species pairs

Of the four focal taxa, only the pied flycatcher and collared flycatcher have overlapping breeding ranges at present (Fig. 2). In sympatric populations of these two species, some hybridization occurs (Qvarnström et al. 2010; Sætre and Sæther 2010). Evidence suggests that some autosomal introgression may occur in these sympatric populations whereas Z-linked introgression is apparently absent (Borge et al. 2005a). This result is consistent with the differential introgression hypothesis. However, heterospecific autosomal alleles have only been found within the hybrid zones and not in adjacent allopatric populations (Borge et al. 2005a). We thus consider it likely that the individuals inferred to possess introgressed alleles in Borge et al. (2005a) are actually recent backcrosses. Wiley et al. (2009) demonstrated that hybrid problems (low fertility) is not restricted to F1-hybrids but also occurs in first- and second-generation backcrosses (and possibly beyond).

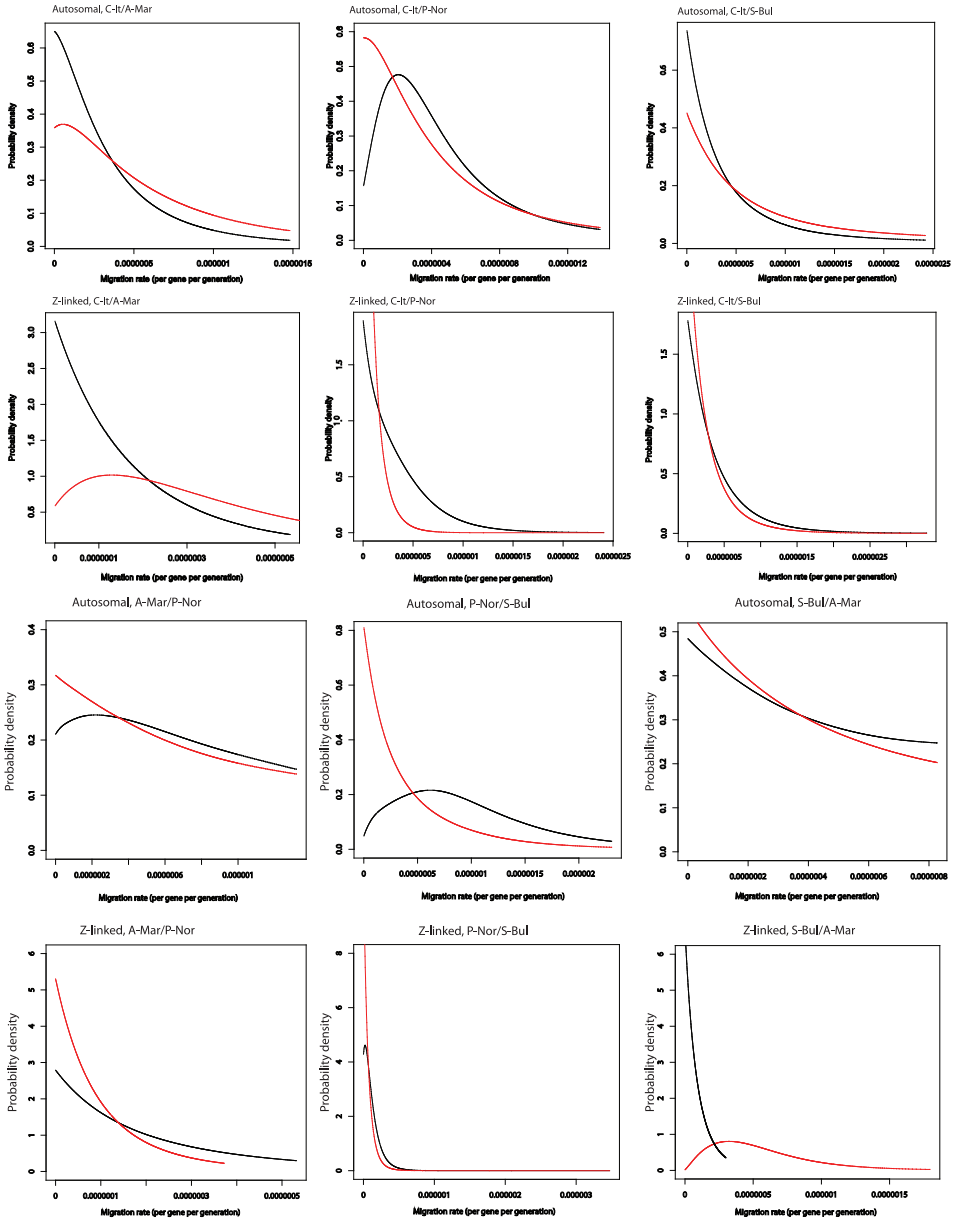


Figure 7. Posterior probability distribution of the migration rate per generation for autosomal and Z-linked loci, respectively. Only values from the run with 10 chains and chain swapping are shown here. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria. Black lines: M1, migration rate from population 2 to 1. Red lines: M2, migration rate from population 1 to 2.

Table 5. Summary of the posterior probability values from IMa-analyses scaled by per gene per generation mutation rate (see methods). The highest point estimate for each parameter is given plus the lower and higher boundaries for the 90% highest posterior density. Only values from the run with 10 chains and chain swapping are given.

Species pair		T[1]	Q1[2]	Q2[3]	M1[4]	M2[5]
A-Mar/P-Nor		430,882	176,067	269,301	2.2	0.067
	A	(105,120–2,227,958)	(64,632–435,339)	(97,691–691,454)	(0.067–114.3)	(0.067–113.0)
		993,988	120,490	248,825	0.027	0.02
	Z	(403,983–2,711,389)	(51,954–251,644)	(123,872–483,418)	(0.027–39.1)	(0.02–22.3)
S-Bul/A-Mar		333,689	234,937	166,468	0.043	0.04
	A	(52,668–1,623,304)	(76,838–657,504)	(51,727–454,920)	(0.043–70.0)	(0.043–70.1)
		1,205,473	351,081	113,618	0.018	32.3
	Z	(530,942–5,800,298)	(189,419–613,046)	(48,085–246,354)	(0.018–19.4)	(3.8–106.5)
C-It/S-Bul		298,086	231,825	136,164	0.12	0.12
	A	(72,478–1,097,840)	(51,694–838,924)	(31,887–416,841)	(0.12–118.7)	(0.1–163.5)
		385,550	225,718	523,381	0.16	0.16
	Z	(201,676–964,166)	(107,441–430,788)	(265,995–1,030,164)	(0.16–96.5)	(0.16–70.9)
C-It/A-Mar		526,701	324,531	111,379	0.074	5.4
	A	(180,796–1,682,073)	(130,989–778,973)	(41,767–285,313)	(0.074–81.6)	(0.074–107.6)
		119,6157	186,969	124,053	0.027	13.0
	Z	(681,592–2,152,259)	(88,101–362,236)	(55,285–255,409)	(0.027–35.5)	(0.080–78.1)
P-Nor/S-Bul		756,159	148,231	191,783	6.2	0.12
	A	(300,063–2,399,3049)	(57,612–338,128)	(81,102–402,084)	(2.2–17.0)	(0.12–10.8)
		1,245,572	311,134	309,255	2.3	0.17
	Z	(627,644–2,310,430)	(172,845–529,287)	(173,330–519,960)	(0.17–2.7)	(0.17–1.7)
C-It/P-Nor		397,431	435,041	106,529	2.0	0.069
	A	(116,150–1,045,707)	(123,672–1,251,683)	(36,384–278,219)	(0.070–8.8)	(0.069–9.0)
		1,096,693	121,179	403,247	1.2	0.12
	Z	(393,300–5,233,845)	(51,195–257,437)	(219,780–707,774)	(1.2–7.8)	(0.12–2.3)

A-Mar = Atlas flycatcher from Morocco; C-it = collared flycatcher from Italy; P-Nor = pied flycatcher from Norway; S-Bul = semicollared flycatcher from Bulgaria.

[1]Time since divergence.

[2]Effective population size 1.

[3]Effective population size 2.

[4]Migration rate from population 2 to 1 (per gene per generation, $\times 10^{-8}$).

[5]Migration rate from population 1 to 2 (per gene per generation, $\times 10^{-8}$).

Hence, current introgression appears to be too low to significantly affect genetic variation of the pied and collared flycatcher except in the very heart of their hybrid zones.

Species distributions are not static and it is certainly possible that some of the species pairs have experienced secondary contact in the past, following their initial split. However, the pattern of elevated divergence and reduced polymorphism at the Z-linked loci appeared rather consistent among the different pairs of species: The Z_θ:A_θ ratio was below the expected ratio of 0.75 in all populations except the Spanish pied flycatcher. However, in a previous study where a larger number of Z-linked and autosomal markers were analyzed, a significantly reduced Z_θ:A_θ ratio was reported also in the latter population (Borge et al. 2005b), suggesting that the apparent heterogeneity in the present study may be mainly due to the lower number of loci included. On the other hand, species differentiation was consistently larger at the Z-linked compared to the autosomal loci. We would think that the degree of historic introgression and the degree to which such introgression would have been biased toward the autosomes are likely to have varied among the species pairs. Accordingly, we would expect the pattern of polymorphism and divergence to have been more heterogeneous than what we observe if differential introgression was the major factor affecting Z-linked and autosomal loci differentially.

IMa analysis

According to the IMa analyses, our estimates of gene flow between the species pairs do not support the differential introgression hypothesis. We found no evidence for elevated rates of autosomal gene flow among any of the species pairs. The estimates of historic gene flow were close to zero for all species pairs at both the Z-linked and the autosomal dataset. Rather, our analyses are consistent with a scenario of classical allopatric speciation. The four focal taxa are well

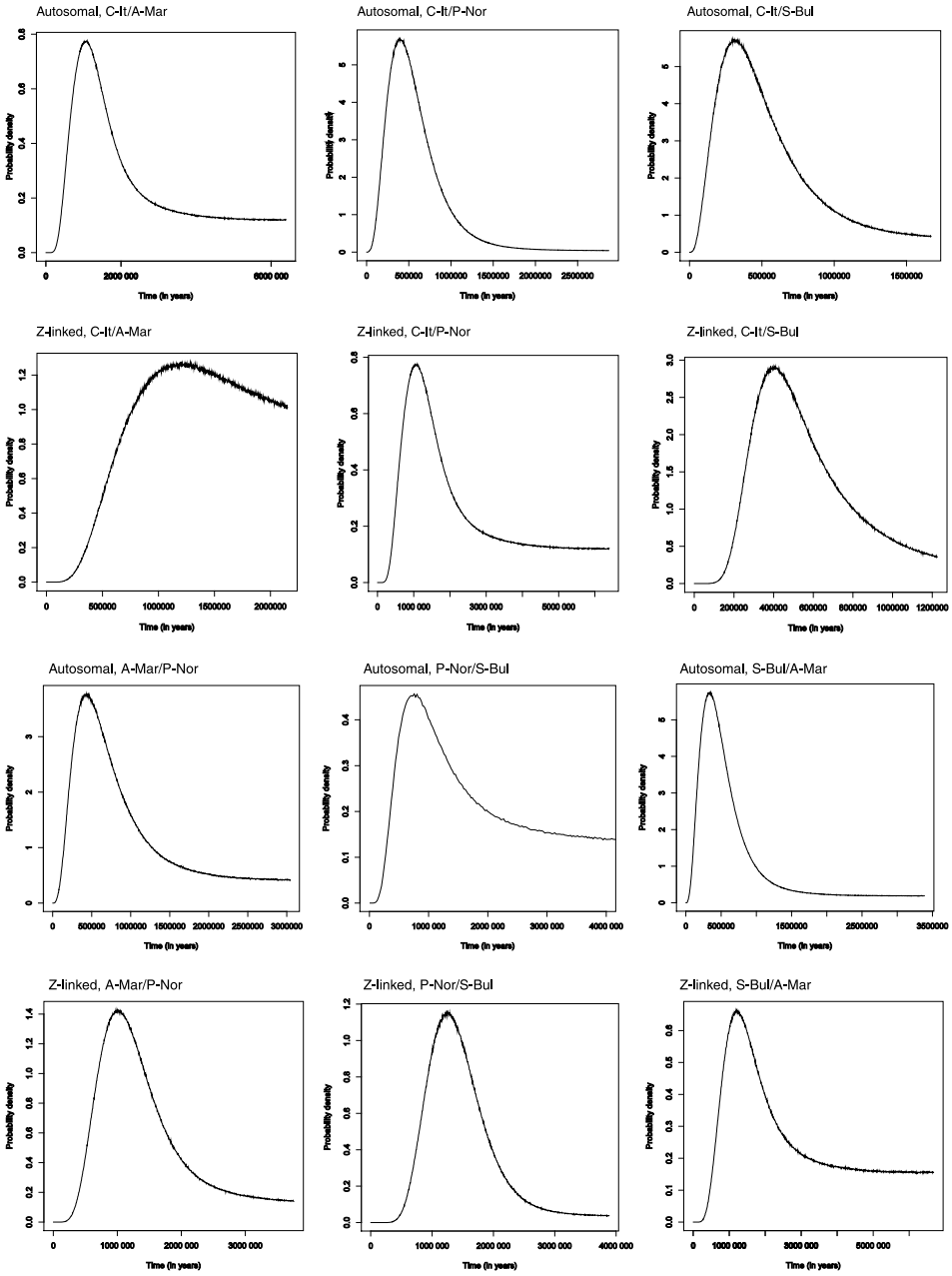


Figure 8. Posterior probability distribution of the time since divergence (in years) for autosomal and Z-linked loci, respectively. Only values from the run with 10 chains and chain swapping are shown here. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria.

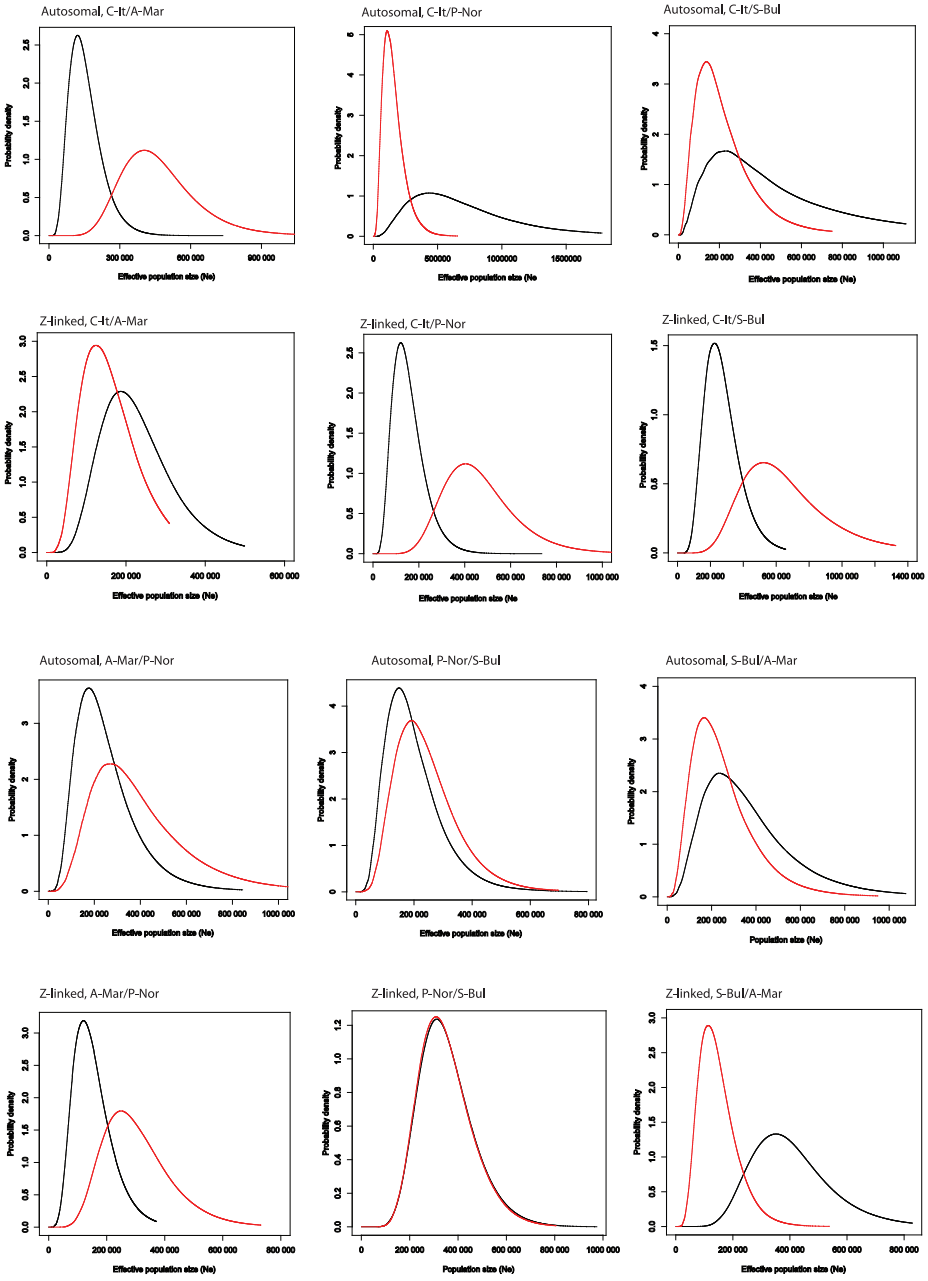


Figure 9. Posterior probability distribution of the effective population size for autosomal and Z-linked loci, respectively. Only values from the run with 10 chains and chain swapping are shown here. A-Mar, Atlas flycatcher from Morocco; C-It, collared flycatcher from Italy; C-Hun, collared flycatcher from Hungary; P-Spa, pied flycatcher from Spain; P-Nor, pied flycatcher from Norway; S-Bul, semicollared flycatcher from Bulgaria. Black lines: Q1, effective population size for population 1. Red lines: Q2, effective population size for population 2.

differentiated according to both the autosomal and the Z-linked dataset, having diverged several hundred thousand years ago or even more than a million years ago according to some of the estimates, not very different from previous estimates based on divergence at mitochondrial DNA (e.g., Sætre et al. 2001).

Faster adaptive divergence on the Z chromosome?

If natural selection has played a significant role in shaping variation and divergence of the Z-chromosome differently from autosomes, one may expect to find molecular footprints of selection events in the data. However, neither the HKA tests nor Tajima's *D* revealed any deviations from neutral expectance, neither at the Z-linked loci nor the autosomal ones.

Mank et al. (2010) analyzed genomic data from the chicken and the zebra finch and found evidence for an elevated d_N/d_S ratio on the Z chromosome. They suggested that genetic drift might be an important contributing factor to this effect. Their argument was that the reduced effective population size of the Z chromosome, reinforced by a female-biased sex ratio due to sexual selection, would elevate the rate at which slightly deleterious nonsynonymous mutations becomes fixed by drift. An operational female-biased sex ratio (some males mate with more than one female) has been observed for both collared and pied flycatchers. In a study by Qvarnström et al. (2003), 4% of the collared flycatcher females were mated with an already mated male, while the corresponding figure in the pied flycatcher is around 10–15% (Lundberg and Alatalo 1992). No such studies have been done on the Atlas or semi-collared flycatchers, but since they are closely related to the pied and collared flycatcher, it is likely that they have a similar mating system, and thus have a somewhat female-biased sex ratio.

Increased rate of genetic drift could explain (or contribute to explain) the observed reduction in polymorphism in our dataset. An elevated mutation rate on the Z-chromosome could in addition account for the elevated rate of divergence that we observe. A higher mutation rate is expected on the Z because mature sperm cells go through more cell divisions than egg cells. Hence, since the Z-chromosome spends two-third of its time in the male germ line, compared to the autosomes one-half, a male-biased mutation rate would elevate the overall mutation rate of the Z chromosome. Indeed, there is some evidence for a male-biased mutation rate in birds (e.g., Axelsson et al. 2004).

Although we acknowledge that genetic drift combined with an elevated mutation rate on the Z chromosome would be consistent with our results, we certainly do not rule out that hemizygous exposure of nonneutral alleles has contributed to the faster-Z effect. Indeed, in a methodologically similar study as the present one, on two of the species included here, and using a larger number of markers, Borge et al. (2005b) reported significant deviations from neutrality among the Z-linked markers according to an HKA test. The latter result is consistent with recurrent selective sweeps on the Z chromosome that would reduce variation within, and increase divergence between the taxa.

Conclusions

To our knowledge, reduced variation coupled with an elevated rate of divergence on Z-linked loci relative to autosomal expectance has been found in all avian cases investigated so far. This consistency suggests that a common evolutionary force or set of forces related to peculiarities of the Z chromosome in itself shapes the pattern. We consider it unlikely that differential introgression is a sufficiently uniform evolutionary force to account for this seemingly general pattern, although it may be a contributing factor in certain cases (see e.g., Carling et al. 2010; Backström et al. 2010). Rather we suggest that the pattern is a manifestation of the faster-Z effect. Further studies are needed to evaluate the relative importance of elevated mutation rates, increased genetic drift, and more effective selection in shaping the Z chromosome differently from the other chromosomes.

Acknowledgments

We thank G. Andersson and G. A. Lien for help in the field, J. A. Anmarkrud for help in the lab, H. Hogner, F. Jakobsen, R. I. Bailey, and E. Rindal for computational and statistical assistance, J. Moreno, J. Török, and L. Z. Garamszegi for providing blood samples, and three anonymous reviewers for valuable comments on previous drafts. Financial support was provided by The Natural History Museum to S.H. and A.J., CEES, University of Oslo (G.P.S., T.B., and S.A.S.), and the Norwegian Research Council to G.P.S.

References

Akey, J. M., M. A. Eberle, M. J. Rieder, C. S. Carlson, M. D. Shriver, D. A. Nickerson, and L. Kruglyak. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. Plos Biol. 2:1591–1599.

Axelsson, E., N. G. C. Smith, H. Sundstrom, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. Mol. Biol. Evol. 21:1538–1547.

Backström, N., and U. Väli. 2011. Sex- and species-biased gene flow in a spotted eagle hybrid zone. BMC Evol. Biol. 11:100.

Backström, N., J. Lindell, Y. Zhang, E. Palkopoulou, A. Qvarnstrom, G. P. Sætre, and H. Ellegren. 2010. A high-density scan of the Z chromosome in *Ficedula* flycatchers erveals candidate loci for diversifying selection. Evolution 64:3461–3475.

Berlin, S., and H. Ellegren. 2004. Chicken W: a genetically uniform chromosome in a highly variable genome. Proc Natl. Acad. Sci. U.S.A. 101:15967–15969.

Borge, T., K. Lindroos, P. Nadvornik, A. C. Syvänen, and G. P. Sætre. 2005a. Amount of introgression in flycatcher hybrid zones reflects regional differences in pre and post-zygotic barriers to gene exchange. J. Evol. Biol. 18:1416–1424.

Borge, T., M. T. Webster, G. Andersson, and G. P. Sætre. 2005b. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. Genetics 171:1861–1873.

Carling, M. D., I. J. Lovette, and R. T. Brumfield. 2010. Historical divergence and gene flom: coalescent analyses of mitochondrial, autosomal and sex-linked loci in *passerina* buntings. Evolution 64:1762–1772.

Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. The relative rates of evolution of sex chromosomes and autosomes. Am Nat. 130:113–135.

Elgvin, T. O., J. S. Hermansen, A. Fijarczyk, T. Bonnet, T. Borge, S. A. Sæther, L. K. Voje, and G. P. Sætre. 2011. Hybrid speciation in sparrows II: a role for sex chromosomes? Mol. Ecol. 20:3823–3837.

Ellegren, H. 2007. Molecular evolutionary genomics of birds. Cytogenet. Genome Res. 117:120–130.

Ellegren, H. 2009. The different levels of genetic diversity in sex chromosomes and autosomes. Trends Genet. 25:278–284.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. 1:47–50.

Fay, J. C., and C.-I. Wu. 2000. Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413.

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. Conserv. Biol. 10:1500–1508.

Haavie, J., G. P. Sætre, and T. Moum. 2000. Discrepancies in population differentiation at microsatellites, mitochondrial DNA and plumage colour in the pied flycatcher – inferring evolutionary processes. Mol. Ecol. 9:1137–1148.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12:101–109.

Harrigan, R. J., M. E. Mazza, and M. D. Sorenson. 2008. Computation vs. cloning: evaluation of two methods for haplotype determination. Mol. Ecol. Res. 8:1239–1248.

Hey, J., and R. Nielsen. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. U.S.A. 104:2785–2790.

Hudson, R. R., and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA-sequences. Genetics 111:147–164.

Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159.

Kreitman, M. 2000. Methods to detect selection in populations with applications to the human. Annu. Rev. Genomics Hum. Genet. 1:539–559.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lundberg, A., and R. V. Alatalo. 1992. The pied flycatcher. Poyser, London.

Mank, J. E., K. Nam, and H. Ellegren. 2010. Faster-Z evolution is predominantly due to genetic drift. Mol. Biol. Evol. 27:661–670.

Nachman, M. W. 1997. Patterns of DNA Variability at X-Linked Loci in Mus domesticus. Genetics 147:1303–1316.

Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics 158:885–896.

Orr, H. A. 1997. Haldane's rule. Annu. Rev. Ecol. Syst. 28:195–218.

Presgraves, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. Trends Genet. 24:336–343.

Primmer, C. R., T. Borge, J. Lindell, and G. P. Sætre. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. Mol. Ecol. 11:603–612.

Qvarnström, A., and R. I. Bailey. 2009. Speciation through evolution of sex-linked genes. Heredity 102:4–15.

Qvarnström, A., B. C. Sheldon, T. Pärt, and L. Gustafsson. 2003. Male ornamentation, timing of breeding, and cost of polygyny in the collared flycatcher. Behav. Ecol. 14:68–73.

Qvarnström, A., A. M. Rice, and H. Ellegren. 2010. Speciation in *Ficedula* flycatchers. Philos. Trans. R. Soc. B 365:1841–1852.

Schlötterer, C. 2003. Hitchhiking mapping – functional genomics from the population genetics perspective. Trends Genet. 19:32–38.

Seutin, G., B. N. White, and P. T. Boag. 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69:82–90.

Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73:1162–1169.

Storchova, R., J. Reif, and M. W. Nachman. 2010. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. Evolution 64:456–471.

Sæther, S. A., G. P. Sætre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, and S. Bures, et al. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. Science 318:95–97

Sætre, G. P., and S. A. Sæther. 2010. Ecology and genetics of speciation in *Ficedula* flycatchers. Mol. Ecol. 19:1091–1106.

Sætre, G. P., T. Borge, J. Lindell, T. Moum, C. R. Primmer, B. C. Sheldon, J. Haavie, A. Johnsen, and H. Ellegren. 2001.

Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. Mol. Ecol. 10:737–749.

Sætre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer and A. C. Syvanen . 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. Proc. R. Soc. Lond. Ser. B 270:53–59.

Tajima, F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596–1599.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldanes rule. Genetics 140:389–402.

Wang, R. L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics 147:1091–1106.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7:256–276.

Weir, B. S., and C. C. Cockerham. 1984. Estimating f-statistics for the analysis of population-structure. Evolution 38:1358–1370.

Wiley, C., A. Qvarnström, G. Andersson, T. Borge, and G.-P Sætre. 2009. Postzygotic isolation over multiple generations of hybrid descendents in a natural hybrid zone: how well do single-generation estimates reflect reproductive isolation? Evolution 63:1731–1739.

Supporting information

Additional Supporting Information may be found online on Wiley Online Library.

Table S1. Fixed and shared polymorphisms between species.
Table S2. Hudson-Kreitmann-Aguade tests.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Table S1. Fixed and shared polymorphisms between species. S-values are the number of variable sites that occur in or between species pairs. S_{shared} indicates the number of shared polymorphisms between the different species; S_{fixed} indicates the number of fixed differences between the species and $S_{average}$ indicates the average pairwise difference between the species pairs. Total-Z is all Z-linked loci combined and total-A is all autosomal loci combined.

Locus	$S_{P\text{-}Spa}$	$S_{C\text{-}It}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	3	1	0	2	3.61	437
BRM-12	12	11	1	1	8.423	1437
CHDZ	3	1	0	1	1.227	638
GHR	4	4	1	1	1.674	555
VLDLR	3	12	0	1	4.514	562
Total-Z	25	29	2	6	19.448	3629
ACLY-16	2	1	0	0	0.133	358
ALAS1-8	6	9	0	0	2.678	290
FAS-Y	1	7	0	1	1.568	551
RHO-1	7	6	0	1	4.333	371
RPL30-3	14	21	9	0	5.582	983
TGFB2-5	3	6	1	0	1.192	402
Total-A	33	50	10	2	15.486	2955

Locus	$S_{P\text{-}Spa}$	$S_{S\text{-}Bul}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	3	1	0	1	1.747	437
BRM-12	12	23	2	2	8.453	1433
CHDZ	3	1	0	1	1.342	638
GHR	4	5	0	1	3.06	555
VLDLR	3	11	0	0	4.362	562
Total-Z	25	41	2	5	18.964	3625
ACLY-16	2	1	0	0	0.133	358
ALAS1-8	6	9	1	0	2.471	288
FAS-Y	1	6	0	0	0.396	551
RHO-1	8	12	2	1	5.613	372
RPL30-3	14	15	3	2	13.2	983
TGFB2-5	3	6	0	0	2.231	401
Total-A	34	49	6	3	24.044	2953

Locus	$S_{P\text{-}Spa}$	$S_{A\text{-}Mar}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	3	0	0	1	1.71	437
BRM-12	12	20	2	2	10.324	1438
CHDZ	3	1	0	0	0.36	638
GHR	4	3	1	1	2.978	548
VLDLR	3	9	0	1	5.381	562
Total-Z	25	33	3	5	20.753	3623
ACLY-16	2	0	0	0	0.1	358
ALAS1-8	6	7	0	0	3.212	290
FAS-Y	1	0	0	0	0.029	551
RHO-1	8	8	1	1	6.54	372
RPL30-3	14	22	9	1	11.382	983
TGFB2-5	3	6	0	0	1.831	402
Total-A	34	43	10	2	23.094	2956

Locus	$S_{P\text{-}Spa}$	$S_{C\text{-}Hun}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	3	1	0	2	3.616	437
BRM-12	12	26	3	1	9.153	1438
CHDZ	3	4	0	1	1.394	638
GHR	4	4	1	1	1.658	555
VLDLR	3	13	0	1	4.621	558
Total-Z	25	48	4	6	20.442	3626
ACLY-16	2	2	0	0	0.163	358
ALAS1-8	6	9	0	0	2.537	290
FAS-Y	1	1	0	1	1.061	550
RHO-1	8	12	1	1	4.812	371
RPL30-3	14	23	4	0	6.208	982
TGFB2-5	3	11	1	0	1.57	402
Total-A	34	58	6	2	16.351	2953

Locus	$S_{P\text{-}Spa}$	$S_{P\text{-}Nor}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	3	1	1	0	0.655	437
BRM-12	12	23	6	0	4.958	1438
CHDZ	3	1	0	0	0.35	638
GHR	4	6	3	0	0.573	555
VLDLR	3	4	0	0	0.581	562
Total-Z	25	35	10	0	7.117	3630
ACLY-16	2	0	0	0	0.1	358
ALAS1-8	6	7	2	0	0.75	290
FAS-Y	1	2	0	0	0.329	551
RHO-1	8	7	6	0	2.388	372
RPL30-3	14	21	11	0	6.964	983
TGFB2-5	3	7	3	0	1.046	402
Total-A	34	44	22	0	11.577	2956

Locus	$S_{C\text{-}It}$	$S_{S\text{-}Bul}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	1	0	3	3.937	437
BRM-12	11	23	0	2	5.756	1433
CHDZ	1	1	0	0	0.181	638
GHR	4	5	0	0	1.663	555
VLDLR	12	11	3	0	2.969	562
Total-Z	29	41	3	5	14.506	3625
ACLY-16	1	1	0	0	0.067	358
ALAS1-8	9	9	1	0	4.084	288
FAS-Y	7	6	0	1	1.905	551
RHO-1	6	11	4	0	3.733	371
RPL30-3	21	15	8	3	12.405	983
TGFB2-5	6	6	2	0	1.9	401
Total-A	50	48	15	4	24.094	2952

Locus	$S_{C\text{-}It}$	$S_{A\text{-}Mar}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	0	0	3	3.9	437
BRM-12	11	20	1	1	8.898	1438
CHDZ	1	1	0	1	1.2	638
GHR	4	3	1	0	2.007	548
VLDLR	12	9	1	0	7.267	562
Total-Z	29	33	3	5	23.272	3623
ACLY-16	1	0	0	0	0.033	358
ALAS1-8	9	7	4	0	3.058	290
FAS-Y	7	0	0	1	1.538	551
RHO-1	6	8	4	0	2.656	371
RPL30-3	21	22	8	0	11.88	983
TGFB2-5	6	6	0	0	1.533	402
Total-A	50	43	16	1	20.698	2955

Locus	$S_{C\text{-}It}$	$S_{C\text{-}Hun}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	1	1	0	0.175	437
BRM-12	26	30	7	0	4.352	1438
CHDZ	1	4	0	0	0.233	638
GHR	4	4	1	0	0.744	555
VLDLR	10	13	8	0	2.898	558
Total-Z	42	52	17	0	8.402	3626
ACLY-16	1	2	0	0	0.096	358
ALAS1-8	9	9	7	0	1.669	290
FAS-Y	7	1	0	0	0.57	550
RHO-1	6	11	4	0	1.8	370
RPL30-3	21	23	9	0	4.915	982
TGFB2-5	6	11	5	0	1.384	402
Total-A	50	57	25	0	10.434	2952

Locus	$S_{C\text{-}It}$	$S_{P\text{-}Nor}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	1	0	2	3.463	437
BRM-12	11	23	2	1	9.451	1437
CHDZ	1	1	0	1	1.19	638
GHR	4	6	1	1	1.8	555
VLDLR	12	4	0	1	4.667	562
Total-Z	29	35	3	6	20.571	3629
ACLY-16	0	1	0	0	0.033	358
ALAS1-8	9	7	2	0	2.515	290
FAS-Y	7	2	0	1	1.838	551
RHO-1	6	6	0	1	4.096	371
RPL30-3	21	21	10	0	9.586	983
TGFB2-5	6	7	1	0	1.634	402
Total-A	49	44	13	2	19.702	2955

Locus	$S_{S\text{-}Bul}$	$S_{A\text{-}Mar}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	0	0	2	2.037	437
BRM-12	23	20	2	0	8.367	1434
CHDZ	1	1	0	1	1.315	638
GHR	5	3	0	0	3.03	548
VLDLR	11	9	1	0	7.068	562
Total-Z	41	33	3	3	21.817	3619
ACLY-16	1	0	0	0	0.033	358
ALAS1-8	9	7	1	0	4.609	288
FAS-Y	6	0	0	0	0.367	551
RHO-1	12	8	4	0	4.493	372
RPL30-3	15	22	5	1	10.797	983
TGFB2-5	6	6	0	0	2.533	401
Total-A	49	43	10	1	22.832	2953

Locus	$S_{S\text{-}Bul}$	$S_{C\text{-}Hun}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	1	0	3	3.943	437
BRM-12	23	26	3	0	6.757	1434
CHDZ	1	4	0	0	0.348	638
GHR	5	4	0	0	1.869	555
VLDLR	11	13	3	0	2.404	558
Total-Z	41	48	6	3	15.321	3622
ACLY-16	1	2	0	0	0.096	358
ALAS1-8	9	9	1	0	3.967	288
FAS-Y	6	1	0	1	1.398	550
RHO-1	12	12	4	0	4.448	371
RPL30-3	15	23	3	5	12.865	982
TGFB2-5	6	11	3	0	2.438	401
Total-A	49	58	11	6	25.212	2950

Locus	$S_{S\text{-}Bul}$	$S_{P\text{-}Nor}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	1	0	1	1.6	437
BRM-12	23	23	3	3	9.486	1433
CHDZ	1	1	0	1	1.304	638
GHR	5	6	0	1	3.15	555
VLDLR	11	4	0	0	4.515	562
Total-Z	41	35	3	6	20.055	3625
ACLY-16	1	0	0	0	0.033	358
ALAS1-8	9	7	0	0	2.41	288
FAS-Y	6	2	0	0	0.667	551
RHO-1	12	7	2	1	5.538	372
RPL30-3	15	21	4	0	12.516	983
TGFB2-5	6	7	1	0	2.656	401
Total-A	49	44	7	1	23.82	2953

Locus	$S_{A\text{-}Mar}$	$S_{C\text{-}Hun}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	0	1	0	3	3.906	437
BRM-12	20	26	2	0	8.926	1439
CHDZ	1	4	0	1	1.367	638
GHR	3	4	1	0	2.063	548
VLDLR	8	13	1	0	6.39	558
Total-Z	32	48	4	4	22.652	3620
ACLY-16	0	2	0	0	0.063	358
ALAS1-8	7	9	3	0	3.238	290
FAS-Y	0	1	0	1	1.031	550
RHO-1	8	12	4	0	3.322	371
RPL30-3	22	23	3	3	12.567	982
TGFB2-5	6	11	0	0	2.129	402
Total-A	43	58	10	4	22.35	2953

Locus	$S_{A\text{-}Mar}$	$S_{P\text{-}Nor}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	0	1	0	1	1.563	437
BRM-12	20	23	3	3	11.558	1438
CHDZ	1	1	0	0	0.323	638
GHR	3	6	1	1	3.106	548
VLDLR	9	4	0	1	5.533	562
Total-Z	33	35	4	6	22.083	3623
ACLY-16	0	0	0	0	0	358
ALAS1-8	7	7	2	0	3.081	290
FAS-Y	0	2	0	0	0.3	551
RHO-1	8	7	1	1	6.515	372
RPL30-3	22	21	12	0	10.19	983
TGFB2-5	6	7	0	0	2.285	402
Total-A	43	44	15	1	22.371	2956

Locus	$S_{C\text{-}Hun}$	$S_{P\text{-}Nor}$	S_{shared}	S_{fixed}	$S_{average}$	Length
Aldob-6	1	1	0	2	3.469	437
BRM-12	20	23	3	3	11.558	1438
CHDZ	4	1	0	1	1.356	638
GHR	4	6	1	1	1.791	555
VLDLR	13	4	0	1	4.773	558
Total-Z	42	35	4	8	22.947	3626
ACLY-16	2	0	0	0	0.063	358
ALAS1-8	9	7	1	0	2.383	290
FAS-Y	1	2	0	1	1.331	550
RHO-1	12	7	1	1	4.568	371
RPL30-3	23	21	3	0	10.288	982
TGFB2-5	11	7	2	0	1.948	402
Total-A	58	44	7	2	20.581	2953

Table S2. Hudson-Kreitmann-Aguade tests. Expected and observed number of segregating sites for each species, and the divergence (average no. of nucleotide differences) between the different species pairs.

	A-Mar		C-It		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	0	1.95	1	1.84	3.9	1.11
BRM_12	20	15.84	11	14.88	8.9	9.17
CHDZ	1	1.28	1	1.2	1.2	0.73
GHR	3	3.57	4	3.4	2.01	2.03
VLDLR	9	11.26	12	10.59	7.27	6.41
ACLY_16	0	0.43	1	0.4	0.03	0.21
ALAS1_8	7	7.85	9	7.38	3.06	3.83
FAS_Y	0	3.49	7	3.28	1.54	1.77
RHO_1	8	6.86	6	6.45	2.66	3.34
RPL30_3	30	25.90	21	24.34	11.88	12.64
TGFB2_5	6	5.57	6	5.24	1.53	2.72
						P= 0.61

	A-Mar		P-Spa		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	0	1.92	3	1.53	1.71	1.23
BRM_12	20	17.69	12	13.24	10.32	11.39
CHDZ	1	1.80	3	1.42	0.36	1.14
GHR	3	4.14	4	3.26	2.98	2.58
VLDLR	9	7.25	3	5.55	5.38	4.58
ACLY_16	0	0.91	2	0.71	0.1	0.48
ALAS1_8	7	6.94	6	5.59	3.21	3.68
FAS_Y	0	0.43	1	0.36	0.03	0.24
RHO_1	8	9.75	8	7.61	6.54	5.18
RPL30_3	22	20.50	14	16	11.38	10.88
TGFB2_5	6	4.64	3	3.73	1.83	2.46
						P= 0.88

	A-Mar		C-Hun		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	0	1.63	1	2.23	3.91	1.05
BRM_12	20	18.13	26	24.83	8.93	11.97
CHDZ	1	1.12	1	1.51	1.37	0.73
GHR	3	2.99	4	4.14	2.06	1.94
VLDLR	8	9.12	13	12.39	6.39	5.87
ACLY_16	0	0.70	2	0.96	0.06	0.39
ALAS1_8	7	6.57	9	8.99	3.24	3.68
FAS_Y	0	0.68	1	0.96	1.03	0.39
RHO_1	8	7.86	11	10.16	3.41	4.39
RPL30_3	22	19.67	23	26.88	12.57	11.01
TGFB2_5	6	6.53	11	8.94	2.13	3.66
						P= 0.84

	A-Mar		P-Nor		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	0	0.95	1	1.01	1.56	0.6
BRM_12	20	20.46	23	20.9	11.56	13.2
CHDZ	1	0.86	1	0.91	0.32	0.55
GHR	3	4.47	6	4.8	3.11	2.84
VLDLR	9	6.92	4	7.22	5.53	4.39
ACLY_16	0	0.00	0	0	0	0
ALAS1_8	7	6.57	7	6.96	3.08	3.55
FAS_Y	0	0.87	2	0.94	0.3	0.49
RHO_1	8	8.27	7	8.77	6.51	4.47
RPL30_3	22	20.74	21	21.25	10.19	11.2
TGFB2_5	6	5.88	7	6.23	2.29	3.17
						P= 0.99

	A-Mar		S-Bul		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	0	1.15	1	1.23	2.53	2.78
BRM_12	20	19.45	23	20.6	2.04	0.66
CHDZ	1	1.26	1	1.34	8.37	11.32
GHR	3	4.16	5	4.49	1.31	0.72
VLDLR	9	10.26	11	10.93	7.07	5.88
ACLY_16	0	0.40	1	0.44	0.03	0.2
ALAS1_8	7	8.69	7	7.99	0.03	0.2
FAS_Y	0	2.40	6	2.73	4.61	3.94
RHO_1	8	9.45	12	10.35	0.37	1.24
RPL30_3	30	21.87	15	12.07	4.49	4.69
TGFB2_5	6	5.62	6	6.14	10.8	10.85
						P= 0.84

	C-It		P-Spa		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	3.43	3	2.54	3.61	1.64
BRM_12	11	14.35	12	10.11	8.42	6.97
CHDZ	1	2.35	3	1.75	1.23	1.13
GHR	4	4.36	4	3.23	1.67	2.08
VLDLR	12	8.88	3	6.4	4.51	4.24
ACLY_16	1	1.46	2	1.07	0.13	0.6
ALAS1_8	9	8.15	6	6.17	2.68	3.36
FAS_Y	7	4.34	1	3.38	1.57	1.85
RHO_1	6	8.08	7	5.93	4.33	3.33
RPL30_3	21	18.91	14	13.87	5.58	7.8
TGFB2_5	6	4.70	3	3.56	1.19	1.94
						P= 0.92

	C-It		C-Hun		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	0.84	1	1.14	0.17	0.25
BRM_12	11	15.91	26	21.64	4.35	4.77
CHDZ	1	2.04	4	2.72	0.23	0.6
GHR	4	3.37	4	4.58	0.74	0.99
VLDLR	10	10.04	13	13.53	2.9	2.94
ACLY_16	1	1.20	2	1.6	0.1	0.35
ALAS1_8	9	7.55	9	10.24	1.67	2.22
FAS_Y	7	3.22	1	4.52	0.57	0.98
RHO_1	6	7.01	10	9	1.75	2.06
RPL30_3	21	18.78	23	25.46	4.92	5.53
TGFB2_5	6	7.05	11	9.57	1.38	2.08

P= 0.99

	C-It		P-Nor		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	2.14	1	2.15	3.46	1.18
BRM_12	11	17.18	23	16.64	9.45	9.62
CHDZ	1	1.25	1	1.25	1.19	0.69
GHR	4	4.62	4	4.64	1.8	2.54
VLDLR	12	8.14	4	8.05	4.67	4.48
ACLY_16	1	0.42	0	0.42	0.03	0.2
ALAS1_8	9	7.47	7	7.5	2.52	3.55
FAS_Y	7	4.31	2	4.41	1.84	2.12
RHO_1	6	6.49	6	6.52	4.1	3.08
RPL30_3	21	21.09	21	20.47	9.59	10.02
TGFB2_5	6	5.90	7	5.93	1.63	2.8

P= 0.85

	C-It		S-Bul		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	2.25	1	2.51	3.94	1.18
BRM_12	11	15.04	23	16.72	5.76	8
CHDZ	1	0.83	1	0.92	0.18	0.43
GHR	4	4.04	5	4.51	1.66	2.12
VLDLR	12	9.83	11	10.99	3.97	5.15
ACLY_16	1	0.79	1	0.91	0.07	0.37
ALAS1_8	9	8.50	9	9.69	4.08	3.89
FAS_Y	7	5.59	6	6.65	1.91	2.67
RHO_1	6	7.95	11	9.13	3.73	3.66
RPL30_3	21	18.86	15	20.83	12.4	8.69
TGFB2_5	6	5.34	6	6.11	1.9	2.45

P= 0.95

	P-Spa		C-Hun		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	3	2.35	1	3.84	3.62	1.43
BRM_12	12	13.99	26	24.08	9.15	9.08
CHDZ	3	2.61	4	4.2	1.39	1.58
GHR	4	2.98	4	4.87	1.66	1.81
VLDLR	0	0.00	0	0	0	0
ACLY_16	2	1.30	2	2.15	0.16	0.71
ALAS1_8	6	5.61	9	8.98	2.54	2.95
FAS_Y	1	0.97	1	1.57	1.06	0.52
RHO_1	8	7.68	11	11.98	4.8	4.15
RPL30_3	14	13.53	23	22.35	6.21	7.33
TGFB2_5	3	4.98	11	7.97	1.57	2.62

P= 0.99

	P-Spa		P-Nor		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	3	1.75	1	2.36	0.66	0.55
BRM_12	12	14.79	23	20.2	4.96	4.97
CHDZ	3	1.63	1	2.2	0.35	0.51
GHR	4	3.97	6	5.35	0.57	1.25
VLDLR	3	2.82	4	3.84	0.58	0.91
ACLY_16	2	0.79	0	1.07	0.1	0.24
ALAS1_8	6	5.25	7	6.91	0.75	1.58
FAS_Y	1	1.28	2	1.66	0.33	0.39
RHO_1	8	6.51	1	8.85	2.39	2.03
RPL30_3	14	15.99	21	21	6.96	4.98
TGFB2_5	3	4.22	7	5.56	1.05	1.27

P= 0.99

	P-Spa		S-Bul		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	3	1.75	1	2.62	1.75	1.37
BRM_12	12	12.78	23	20.01	8.45	10.66
CHDZ	3	1.63	1	2.44	1.34	1.27
GHR	4	3.67	5	5.51	3.06	2.88
VLDLR	3	5.49	11	8.45	4.36	4.42
ACLY_16	2	0.97	1	1.51	0.13	0.65
ALAS1_8	6	5.56	9	8.31	2.47	3.6
FAS_Y	1	2.32	6	3.54	0.4	1.53
RHO_1	8	7.94	12	12.33	5.61	5.34
RPL30_3	14	13.32	15	19.92	13.2	8.96
TGFB2_5	3	3.56	6	5.35	2.23	2.32

P= 0.96

	C-Hun		P-Nor		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	2.50	1	1.9	3.47	1.06
BRM_12	26	27.35	23	20.07	10.26	11.84
CHDZ	4	2.88	1	2.26	1.36	1.25
GHR	4	5.39	6	4.11	1.79	2.3
VLDLR	13	9.98	4	7.54	4.77	4.25
ACLY_16	2	0.97	0	0.74	0.06	0.36
ALAS1_8	9	8.60	7	6.56	2.38	3.22
FAS_Y	1	2.04	2	1.53	1.33	0.76
RHO_1	11	10.24	7	8.26	4.54	4.04
RPL30_3	23	25.71	21	18.95	10.29	9.62
TGFB2_5	11	9.34	7	7.12	1.95	3.49
						P= 0.90

	C-Hun		S-Bul		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	2.65	1	2.24	3.94	1.5
BRM_12	26	24.85	23	20.88	6.76	10.02
CHDZ	4	2.36	1	2.03	0.5	0.96
GHR	4	4.46	5	3.77	1	1.77
VLDLR	13	11.73	11	10	2.4	4.67
ACLY_16	2	1.39	1	1.21	0.1	0.5
ALAS1_8	9	9.90	9	8.55	3.97	3.51
FAS_Y	1	3.77	6	3.28	1.4	1.35
RHO_1	11	11.99	12	11.01	4.51	4.51
RPL30_3	23	23.16	15	19.43	12.87	8.27
TGFB2_5	11	8.74	6	7.59	2.44	3.11
						P= 0.75

	P-Nor		S-Bul		Divergence	
Locus	Observed	Expected	Observed	Expected	Observed	Expected
ALDOB_6	1	1.29	1	1.49	1.6	0.82
BRM_12	23	19.81	24	23.53	9.49	13.15
CHDZ	1	1.19	1	1.37	1.3	0.75
GHR	4	4.37	5	5.03	3.15	2.76
VLDLR	4	7.58	11	9.03	6.39	4.78
ACLY_16	0	0.38	1	0.45	0.03	0.21
ALAS1_8	7	6.78	9	7.98	2.41	3.65
FAS_Y	2	3.15	6	3.79	0.67	1.73
RHO_1	7	9.00	12	10.66	5.54	4.88
RPL30_3	21	17.70	15	20.89	12.52	9.92
TGFB2_5	7	5.57	6	6.79	2.66	3.11
						P= 0.99

A-Mar = Atlas from Morocco, C-It = collared from Italy, P-Spa = pied from Spain, C-Hun = collared from Hungary, P-Nor = pied from Norway, S-Bul = semicollared from Bulgaria